1/39
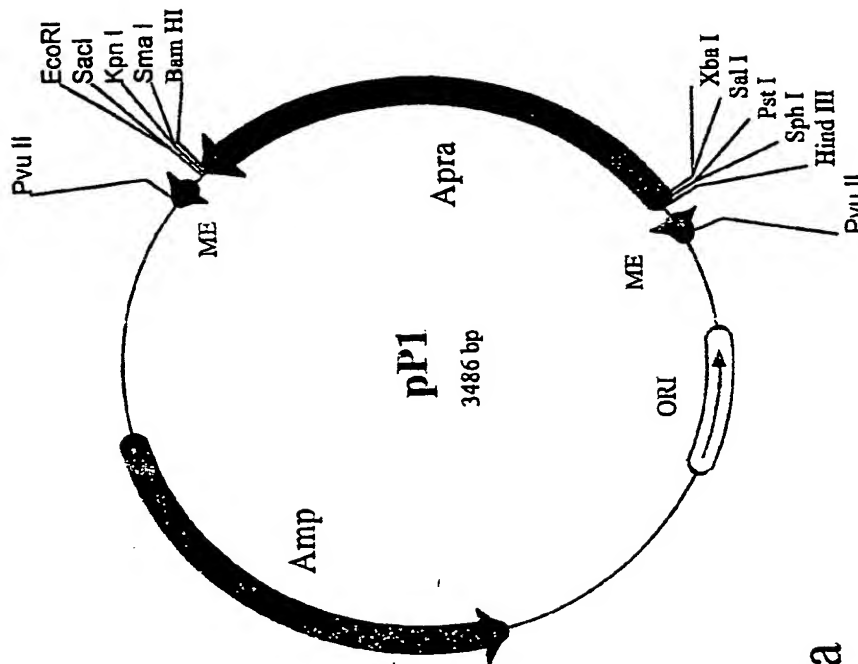

Figure 1a

2/39
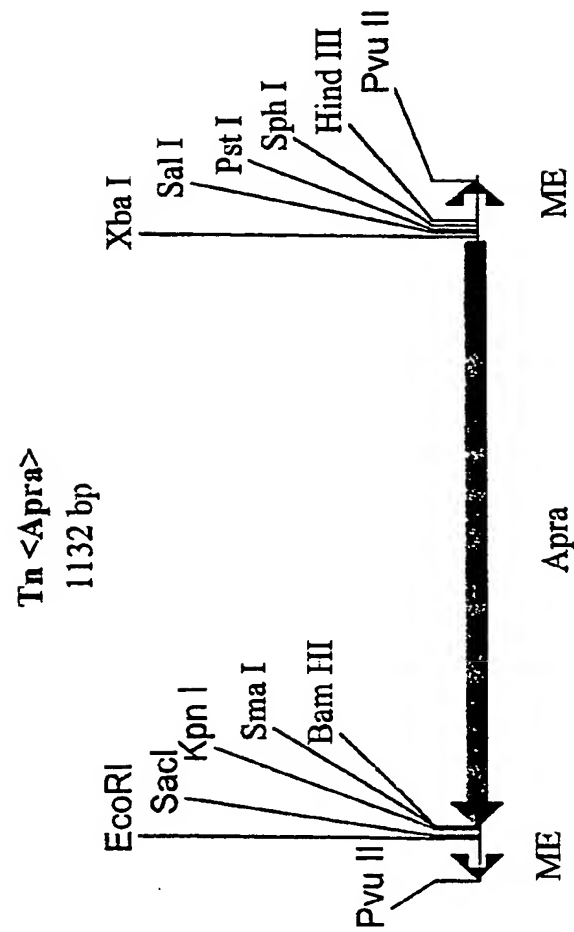

Figure 1b
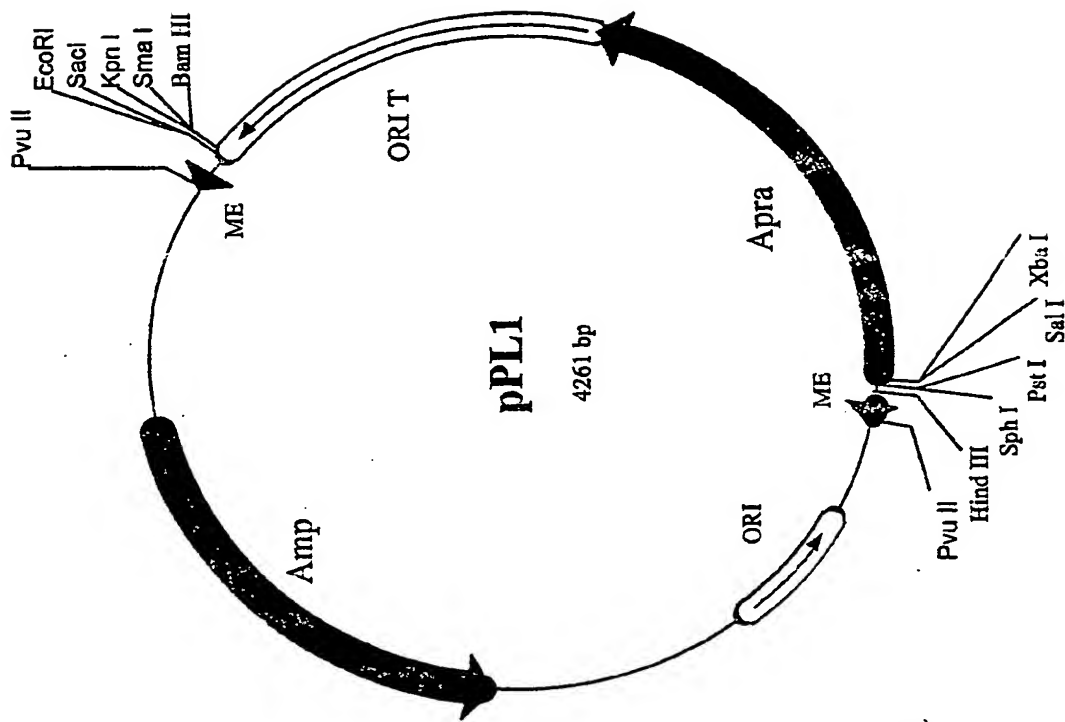
3/39

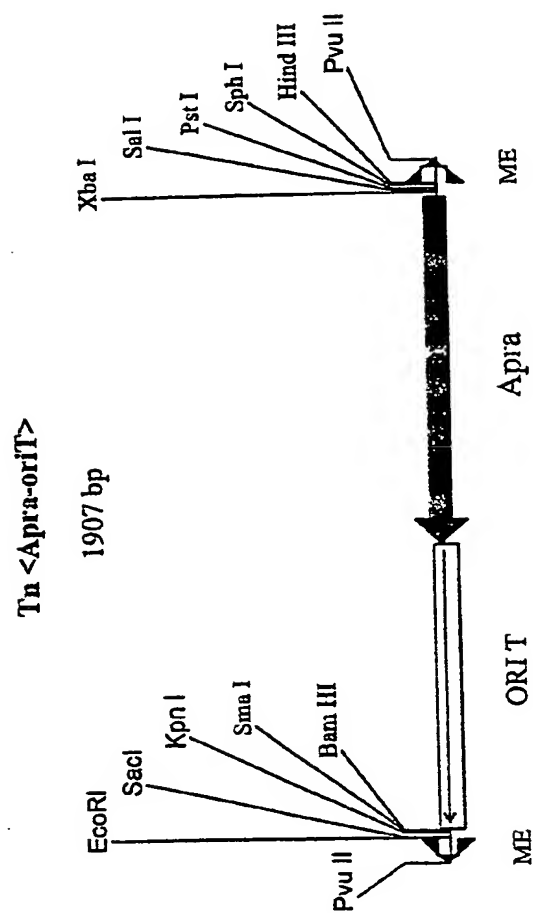Figure 2a

Figure 2b

Figure 3a

Tn <Apra-oriT-att-int>
3924 bp

Figure 3b

Figure 4a

10/522037

cloned insert intact

A gene

cloning vector

Apra
ORI T
integrase

X gene inactive
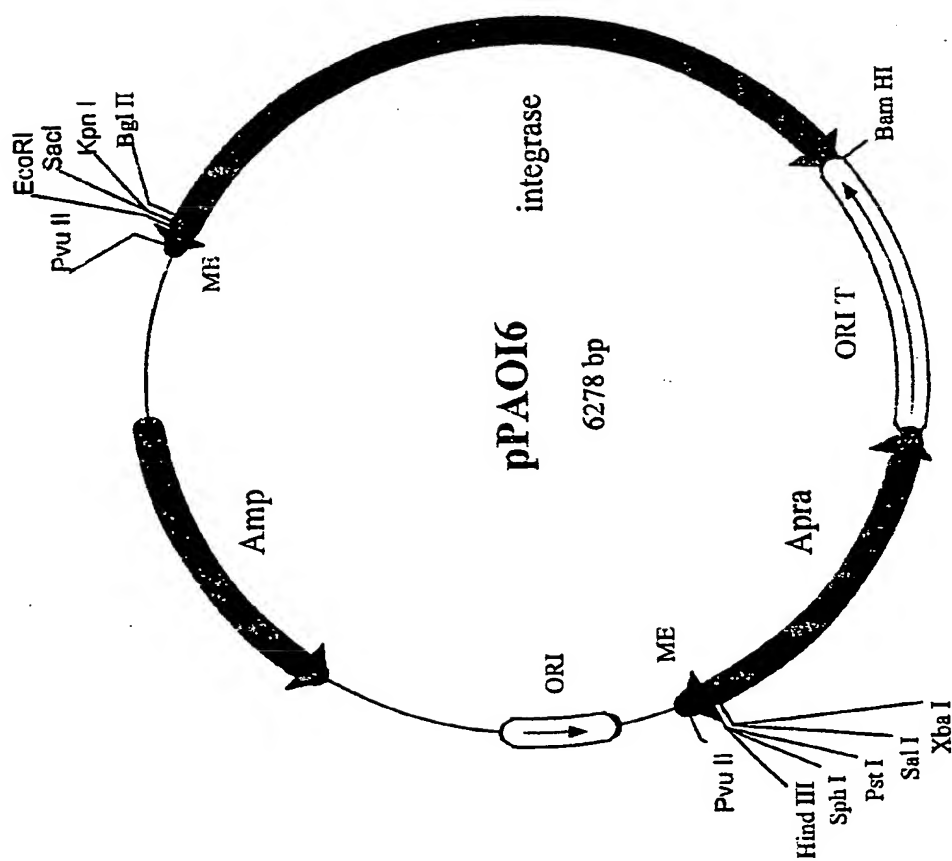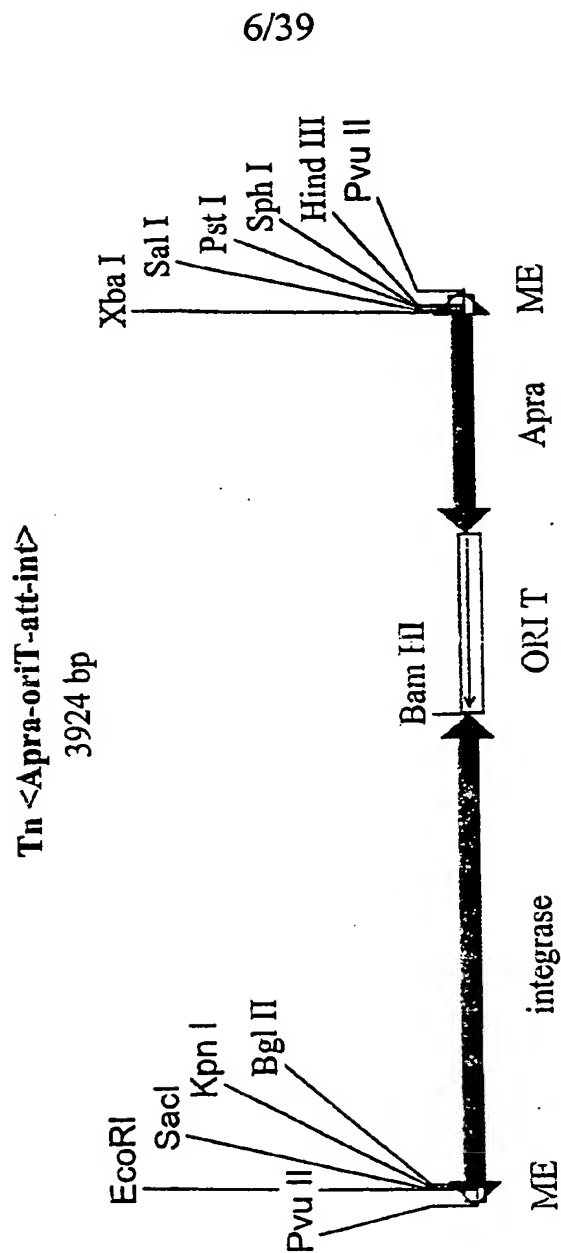
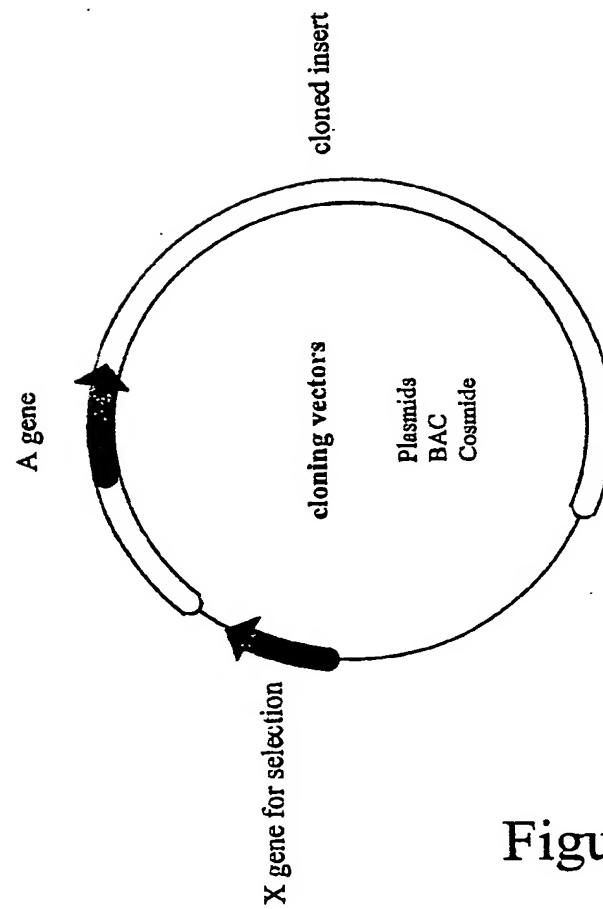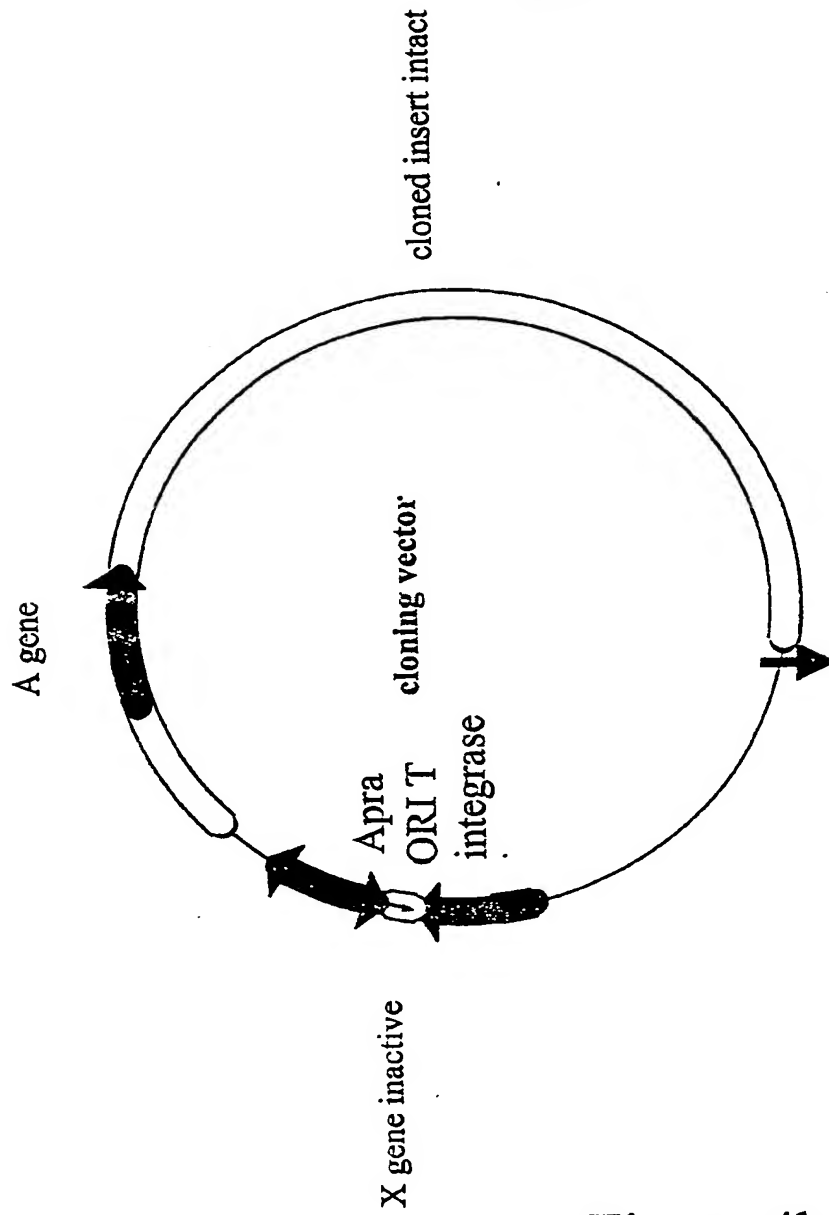Conjugative and integrative plasmid

Figure 4b

10/522037

Figure 4c

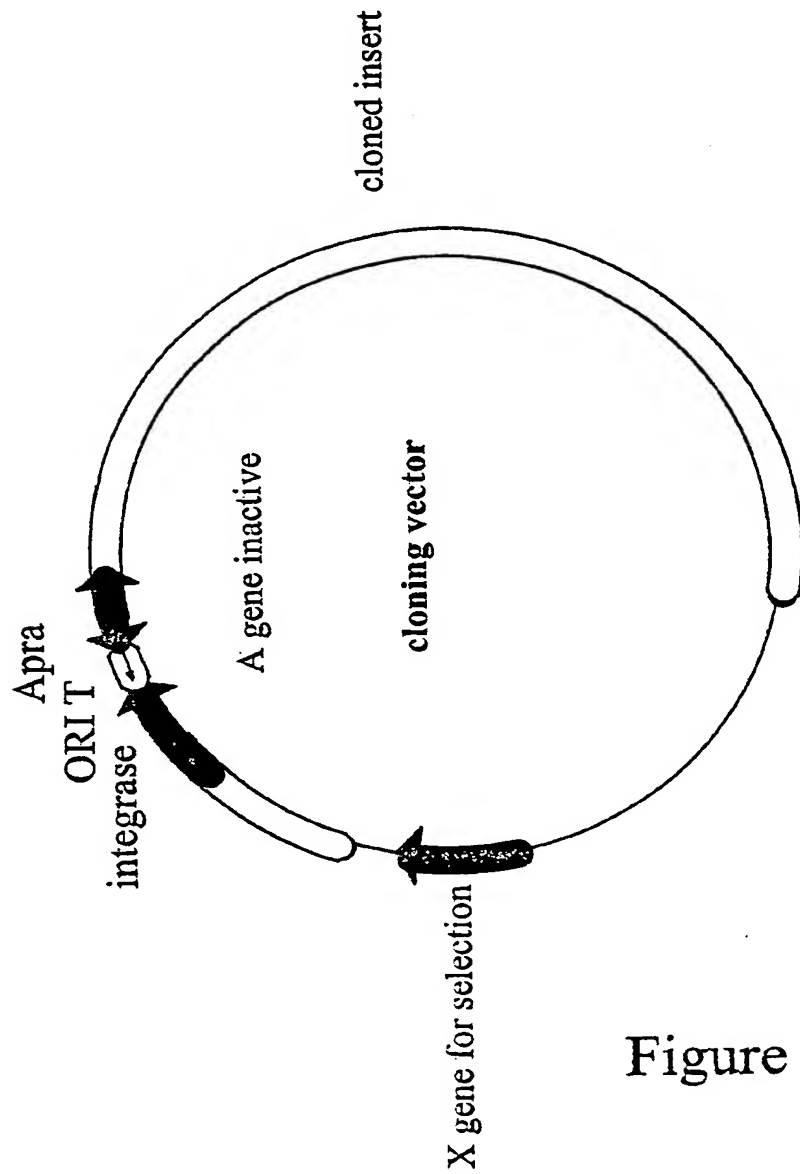Figure 5a

```
CDS              76..1134
                 /note="ABC_transportr"
                 /gene="TAP2 PROTEIN"
                 /blastp_match="Oryzias latipes"
                 /blast_score= 0.002
CDS              complement(1096..2430)
                 /note="none"
                 /blastp_match="Anabaena sp"
                 /gene="ALR1117 PROTEIN"
                 /blast_score=2e-18
CDS              1178..1624
                 /note="Gram_pos_anchor"
                 /gene="CELL WALL SURFACE ANCHOR"
                 /blast_score=1e-04
                 /blastp_match="Streptococcus pneumoniae"
CDS              complement(2506..3567)
                 /note="CONSERVED"
                 /gene="HYPOTHETICAL PROTEIN"
                 /blast_score=0.019
                 /blastp_match="Deinococcus radiodurans"
CDS              complement(2906..4222)
                 /note="glycosyl transferase"
                 /gene="lipopolysaccharide"
                 /blast_score=2e-23
CDS              complement(4092..5321)
                 /note="glycosyl transferase"
                 /gene="glycosyl transferase"
                 /blast_score=1e-15
CDS              complement(6337..8502)
                 /note="PUTATIVE"
                 /gene="GLUTAMINE AMIDOTRANSFERASE"
                 /blastp_match=" Bordetella bronchiseptica"
                 /blast_score=1e-16
CDS              complement(8181..9530)
                 /note="none"
                 /gene="MEMBRANE PROTEIN"
                 /blast_score=0.035
CDS              complement(9531..10721)
                 /note="NOEC Transmembrane"
                 /gene="NODULATION PROTEIN"
                 /blastp_match="Azorhizobium caulinodans"
                 /blast_score=3e-07
CDS              complement(10504..11274)
                 /note="PUTATIVE"
                 /gene="HYDROLASE"
                 /blastp_match="Streptomyces coelicolor"
                 /blast_score=4e-14
CDS              12874..13689
                 /note="HYPOTHETICAL Meth-transf"
                 /gene="PROTEIN PA1088"
                 /blast_score=2e-06
                 /blastp_match=" Pseudomonas aeruginosa"
CDS              14195..15976
                 /note="PUTATIVE, Glyco_transf"
                 /note=""
                 /gene="LIPOPOLYSACCHARIDE BIOSYNTHESIS"
                 /blast_score=2e-06
                 /blastp_match="Vibrio cholerae"
CDS              15427..16512
                 /note="PATHWAY: INNER CORE LIPOPOLYSACCHARIDE
                 BIOSYNTHESIS"
                 /gene="PHOSPHOHEPTOSE ISOMERASE "
                 /blast_score=3e-17
                 /blastp_match="Helicobacter pylori"
CDS              15579..16253
```

```
                        /note="none"
                        /gene="PHOSPHOHEPTOSE ISOMERASE"
                        /blast_score=2e-22
                        /blastp_match="Neisseria meningitidis"
        CDS             complement(16505..17656)
                        /note="BIOSYNTHESIS PUTATIVE"
                        /gene="LIPOPOLYSACCHARIDE"
                        /blast_score=2e-17
                        /blastp_match="Thermotoga maritima"
                        /pfam_match="Glycos_transf_1"
        CDS             complement(17657..18697)
                        /note="none"
                        /gene="ALR3073 PROTEIN"
                        /blast_score=6e-27
                        /blastp_match="Anabaena sp"
        CDS             complement(18615..19304)
                        /note="none"
                        /gene="ALR4487 PROTEIN"
                        /blast_score=8e-07
                        /blastp_match="Anabaena sp"
        CDS             complement(19301..20596)
                        /note="ATP_GTP_A"
                        /gene="ABC TRANSPORTER"
                        /blast_score=3e-61
                        /blastp_match="Synechocystis sp"
        CDS             complement(20535..21476)
                        /note="PERMEASE COMPONENT"
                        /gene="POLYSACCHARIDE ABC TRANSPORTER"
                        /blast_score=6e-41
                        /blastp_match="Clostridium acetobutylicum"
        CDS             complement(22025..22951)
                        /note="involved in the synthesis of a polysaccharide
                        capsule ?"
                        /gene="32.3 KDA PROTEIN"
                        /blast_score=3e-17
                        /blastp_match="Sphingomonas sp"
        CDS             23155..26523
                        /note="peptide syntase"
                        /gene="mcyA, mcyB and mcyC"
                        /blastp_match="Microcystis aeruginosa"
                        /blast_score=0.0
        CDS             26409..34433
                        /note="polyketide syntase et peptide syntase"
                        /gene="mcyD, mcyE, mcyF and mcyG"
                        /blastp_match="Microcystis aeruginosa"
        CDS             34418..37500
                        /note="CYSTATIN"
                        /gene="PEPTIDE SYNTHETASE"
                        /blast_score=0.0
                        /blastp_match="Anabaena sp"
        CDS             35359..37500
                        /note="gene cluster"
                        /gene="nostopeptolide biosynthetic"
                        /blast_score=2e-41
                        /blastp_match="Nostoc sp"
      Sequence 37500 BP; 6199 A; 12698 C; 12769 G; 5834 T; 0 other;

   gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa agggggatgt gctgcaaggc        60
   gattaagttg ggtaacgcca gggtttttccc agtcacgacg ttgtaaaacg acggccagtg       120
   aattgtaata cgactcacta tagggcgaat tcgagctcgg tacccgggga tcccacgtac       180
   cacgagctca tctggaagag cccgctccag acacccgagg aatggaagcc ggcgtcgtg        240
   agaccgtaca cgcagaagcg tctcgtcgcc ttcgttgact ggcttgcgac gacgatgaag       300
   ctggatctca ccaggatgtt cgccgcaggc agctcgatgg gcggatcggg cgcgatcatg       360
   ctcgcgattc gctatcccgc acgattcgcg tggaccgtgt cgtgggtcgg cgtccacgtg       420
   cccgccgact ctccgcaatt cacatcgtcg tacgagctgg tgtacggccg gcccgacttg       480
   aaggtgccgt tcgagaacgg cacgccggtc tgggatcatt ttagtgacgt ctggtacctg       540
   cggcagcatc cggagcagga catcgggttc atcacgttct cgaacggcaa gaacgactcg       600
   gcgatcggct ggcgccaggc cgtcgaattc ctgaagacgc tgcaggagac caggcaaccg       660
   cacctgttcg tctggggcca ggagggacac ggccagcgcg cgaagatgcc ggaaggtggc       720
```

```
ggcgaacggg agatgccgct cgacctcaga acgaaccaga gcctgccggc gttcagccgt    780
tgcacgctcg acgacgatcc gggagacggg tcggattcga gcggcgtccc ggccgggcag    840
atcaactcct acgtcacctg gcagcctgac tcgatcgtcg acgaacccgg ccgatggagc    900
gtcgtcatca agctggcgaa tcgaccgccg cggtcgcccg ccgcagtcga cgtcacgccg    960
cggcggctgc agcagttcaa gctcaagccc ggcgatcagg tgacatggac gaactccgcg   1020
ggcaacaccg tcgtgcagcg cggcgaagcg gtcgcggacc ggtggggtct cgtcacgctc   1080
ccgcagatgc aggtctcaac gggtggaaac aggatcctgg tgagcaggaa ctaggcgggc   1140
ggcgtcctcg acgagccggc gcatcggcgc cagttgatgt gcccacgaaa gctggctgag   1200
gtccgcgcgc gcggcgatgt cccgtcgcgc ctgcggcggc tttgccagca aatccagcac   1260
agccgcccgg aacgcctcga ctgattccgc cacggagcac gccggccgtg cgatggtggg   1320
aagtcccccct gcgacggccg gggtcacgac gttcgggagc cccgcggcaa tcgcttcgag   1380
cactttgttc tgcacgcctc tcgcgatcgc aatcggcgcg accgcgatcg ccgatcgcca   1440
gagatactcg cgtacctccg ggaccgcgcc ggtcacttca atcgtgggat cggccgcgag   1500
tgcccgcacc gcggcggtcg ggtccatgcc cacgagcatc agccgcgcgt ccgggcgggc   1560
gcgaaccacg gacggccaga cctctcgcgc gaaccagagc gccgcgcgct cgttgggctc   1620
gtagccgaac acgccgcaga agacgacccg cgcttcggcg gccggtggtc ctttcggcgc   1680
aaacatctcg acgtcgatgc cgttcggcac gaccgacgcc gtgaggccgc ccgtgacttc   1740
ttcgagcagc aggcgctcgc gctcgctcac gaccgtggtc gccgccgcgc gccgcatcgc   1800
gactttctcg aagcggagca agcggccggc ctcgcggcgg aacagccagc ctttcggcca   1860
cgccgccgtc gacgccagcg cgcgccattt ctcggagtcg acgtccacca tgtccaggac   1920
gaagggatc ccgcccagga tcggctcgca ggcgtagcgg gccatcccgc tgcagtacgc   1980
gagcaccgcg tcgggccgcg tcgcggcaat ccgtcgttcg agcacctggt agatctccgg   2040
cgaatgcagc aggacgtgcg tgaggggtcg atcgcccggc agcgccagac cgctcgcgat   2100
gagattccgg ccgcggcgga cgcgcacgac gtcggtcgac gccgtcacgc cggcgagatc   2160
cgctcgatgc gaccactcct cgtcgtcgtg cgcgagcgag acgaggtgca cgtcggcgga   2220
tcgtgcgagc gtgtgaatga gatgaaacgc gcgaatgcga tcgccccggt tcggggcgta   2280
gggcagccgg tgtgtgagca gaaggactct caccgtgcgg catccgatcg ctctcggatg   2340
acgcgggccg gtacgcccgc ggcgatcacg cgggcgggaa tcggccggt cacgaccgac    2400
cccgccgcaa tcacggaatc cggaccgacg tcagccatga ccacggcggc gctgccgatc   2460
caggtgccgg cgccgatccg caccatccgc ggtgtgcctg gctgctcacg gatcggccgc   2520
gtgaggtcgc tcgtgccgtg cgtgtccggt ccgctgggga tgtggacgcc ggcgccgacc   2580
aggacatcgc gctccaggtg gacccagccg agatggcatc cgggtcccac gtacacatgc   2640
tcgtcgagct tcgcgccggt ctgcgaaaag atggtgccga aggcgaccgt gaccgacggg   2700
tcgcagtgtg caagcacgcg gccgaggaaa gccgcgcgaa ggtattgtcc gatgacgccg   2760
gggatgagcg agagccactg cgtcgacccct tcgagggccc ggtcggcgcc gaggaaggge   2820
cgccgcacgg agtaggagag caaggcgggc aggacgagca ccagcgcgat gccacgggcc   2880
agtcccttgc cggcgtcctt gaggctcacg cccactgcag cgcgcgcccc gcgtcggacc   2940
gggctccgta tcgtgtgacg agcgagtcgt agagcgcttc cactttgcgc atccgccgct   3000
cgaacgataa ctcttcttcg atcggcgtc gcgccgcaat cgcacgggtg cgtccggccg    3060
cgggattggc catcgcctgc tcgatcgcgt ccgcgaggcc tgacgcccgg ccgggctgaa   3120
tgaccagggc gtgcgtgccg ttcgacacga gctcggccgt gcccccggca gcggtcgcga   3180
cgatgggtgt ctcgaacgcc atggcttcga gcacggcgtt gggcgtccct tcgtagtccg   3240
acgactgcac gagcaggtcc agcccgtgat ggacggccga gacgtcggac gtgtggccga   3300
ggaaacgcca ggttccgggc ccgagctcgc gggcggcctg ttcctcgagc ggacgccgga   3360
ggctgccgtc gccggcgatg acgagacgga ggcgcggcca cttcgcctga agggcggcac   3420
aggcttcgat gagcaagtcg aatcgtttct gcggctcgag acggccaacc gcgccgatga   3480
cgaagtcgct ggcctcgaat cccagtcctt cacgcgccgc ggcctcagat cgtggatcgc   3540
gtcggaaagc atggtggtcg atgccattcg ggatcgtgac gaccgatcc ggccgtccgc    3600
ctttggccag gagctcccgc cgaatctgtt cggacacggc aatgacggca gggaagcggg   3660
cgagaatccg gcgatcgagc gggtagtaga cggtgcgttc gcgaaacgaa tggccggtcc   3720
atccatgggc cgtggccagc ggaatcacgc gctcggcgcg tcccaggagc agggcgagga   3780
gatcggtctt gtactcgtgc gcgtgcacga tgtcgatgcg cagatctttc acgagcgcgc   3840
ggagcttcga ccagattccg gtgtcgaacg aatgcttctc ccggacctcg acgtagtcga   3900
tgccggccgc tttcgcccgt tcgtccatgc cgaacaccgg atcgcgctgg tcgcggatgt   3960
agcagacggt gatcgcgtac ttggagcgat cggcctgcgc cgttccgagc agaatcgtct   4020
tctccggtcc gccgccgtg cccgcacgc tgcggagctc gagcacgcgc acggacgcg     4080
cgattccgg cctaggcgcg gcgtctggca gcatggcccc ccggccgac cagcgcgatg     4140
gcccgggcga tatcgtacgt gccggccacg acgagcgact gtgcgatctc gacgagtccc   4200
actcgccgc gccggcgccg gatcatcgcg gctgcgcgca agccggcgaa accggcaagc    4260
agcaggagcg cggtaccgac aatccaccac ccgccgcgct gccacagcgc gaggccagc    4320
ggaatcgaga cgagcgccag cagatcgagg accgggatga tcacgctggg cagatcccgc   4380
atggcgagcc ggtgacggaa gctcacgcgc aggttgtcgc ggccggcca cagctcgccg    4440
aaaaacacgg cgcgaagcga tcgggcgcagc ccctggtgca cgctgcggag ccgcgggtcc   4500
gacaggatgc ggccgcgcc ccggcgcgg tcgaaggcct cgcgcttcac ggcaagattg     4560
tcgaggctgg tatcgaagcc accggcgcgg tcgaaggcct cgcgcttcac ggcaagattg    4620
ccgctcggga gccagtccac gtcttcggtg ctctgcccac gacggcgaaa cgagtcgtac   4680
gcgcgctgca cccagttgcc gtcaccggga gcttcgtact gcgcgccggc agcgacgacc   4740
cccgggtct ggagcagatc gacggcgacc accggccagc gcgggtcgat ctcatgatcg    4800
gcgtcaacga acgccagcac gtcgccctcc gccgaccggg ccgcgcggtt gcgaagcgtc   4860
```

```
gccacgttca cgtcgggcag cgacagcact cgcgcggatc cgcgctctgc cacggcacgc    4920
gagtcgtcgg tcgacccgtt gtccgccacg acgatctcgt accagccagc cggcgcctcg    4980
gtgcgatgaa tcgacgccag gcaacggctg aggtgcgccg cgccgttctt gacggggatg    5040
acgaacgaca cgcgcggcgc cacggcaatt ttcgcggtca tcgatttcgt agcgcggccg    5100
gctttcgtag cgcggcccgg gttggggggcc gcgagaactc gaacgcgccc agatccggcg    5160
cccctgctct cggccgccga tcgaaatcga gcggcgcctg cggaatctcg atgccggcgt    5220
cgatcgcctc cctggcgggct gccgacaggt gcagatcgca cgaggccgcg tttgcgaacc    5280
actctgcggc ggcgttcgtg acgttgttcg tcgcctccgc ctcaccgccg ttacggcgca    5340
ggatctgccg gctcgtcagg ttgttgcgga tgagcccgtt cgtcacggga aaccggacgc    5400
cgatcgtcca cggcgtcgcc gtgccgagcg tgaagacgct gttgtgctcg atccggaagt    5460
cggtggacgc gttggcttcg atcgcttcgt cggcccattc gttcaggttg cagatcacgt    5520
tgttccgcac gatgccgcgc tggtgatcga tctgcttctc gccgtcgcgt gcgtaggccg    5580
tgatccccgc ggccaggccg atggcgatgc ccctgaacga atcgacgatg acgttccgtt    5640
cgatgagcgt gtcctgcgag ttcgcccagg cgaggatcgc cgggcgggat cgccagccgc    5700
ccgattgcgg gccgcgaatc cgcaggaaga cgttgtcgcg gatgacccac ccccgggccg    5760
ccaggatgtc cacgccgtcg gtgtaatcgg acggcgcgct cgtcgtgtac tcgaagcgcg    5820
agcacgccac gaggccatcg tccgcgaagc gtccgtcagt gccgacgctg cctttgagca    5880
attgctgtcc cgcatcgatg agctggacgt tgtgcaccgt ggctcgggac gcgccaagct    5940
cgccgcgcac ctgaatcgcg tggtacccga cgtggccgac cgtcagatcc gcaatcgtca    6000
cgtcagtggc gccaaccgag agcgccacgc cgatcgcgtc gccggtcatg ccgccgccgc    6060
gaatgacgac gcggtcccga tcgccggacg ggctccggag cacggtgccg gggacgttga    6120
tgctggccat ccggtcgagc cggtactcgc cgttctcgag cagaatcgtg gtgcgcggct    6180
tgacgcgttc gagcgcctgc agcagctcgg acgtccgccg cacggtgacg acctgtttcg    6240
tcgccggcgc cgcccatccg cagaagtgcg gcaccggcgc ggcggcctgg ccggacggca    6300
gaagcgccac gagtgcgccg gccacgatga ggcgagtcat gccgccaccg tgtcgacgtc    6360
gtgttcgcgc aaccagagct cgaggatgag caggatttga aggaggtagc cgcgcgactg    6420
tctcccgtga cgcgtgtcgt cgatgatggc gcgcaacgtc tcggggcgca ggaacatccg    6480
cgtgcgcgcc tccggggcga acagctcgcg ctcgacggct tcgagcagtc gcgtctgcag    6540
ccagcgatcg aagtcgtgat agtgacggta cccaggcgcg ttcaggcgct tcagcagggt    6600
gttgaccttg tcgaacacct tctccgccag cggaccggcg tcgccgcgcg cgccggtgtt    6660
cgagttccgc accttgagca gcgccgggtt gccggcctgc gtgatcgcgc gatgaagcgc    6720
cgtgtcgtcg cgccatgccg gatcgccgct cagcaggacc tgcaggaact cgcgatccat    6780
gaacggcagc cgcacgtcca cgacggtgcg gaagagatcg agcgagggaa tcgtgaagcg    6840
ccgatgatgc tcacgaagat agagatagct gctgagctcc gcgggtgaga gctccacggt    6900
gtcgaggagc gcgtgcagcg agtcggccga gccccgtccg gagaggcgcg cagcgtctgc    6960
tgtgaagagc gtgtccagcg gcaggccgct gctgatgtag ttggcgcgcg acgtcatgta    7020
gtcgatgaac tcgccgcggc ccgtcatcgc gaacacgtgg ccgtcggtgt ggagcggcca    7080
cgccaggctc gtcttggcga gctctccgcc gtggccgcgc agcaggacct cgataccgga    7140
gtcccgcagg aaatcgatcg ccagcatctc ggtcagtccg tggctcagat acatgccatc    7200
ggtgagcgag accatgcgcg cctggttcgg cagaaagtcc ttgaggtacc ggtcgtcgag    7260
ctcgaagaag cggtggtcgg taccggtgag cttcgacagc tgttgcgtga tcacctgatc    7320
ggcgcagccg cgcgacgccca cgcgtgtacgt gatgagcccg ccaacaccgt tcaccgcact    7380
cagcacggcg cggctgtcga ggccgccgga gagcgacagc ccgaagcgat gcctgcccga    7440
gagcgacttc tcgaccgcgc gcctgaacgt cgtgtgcacc gcgtggacgt acgcgtcctt    7500
cgtcgtacgc ggtccgcggt agagatcggc cggctcgaaa taccgggaca cgctcgcggc    7560
gcctgacggc cagtcgtacg cgagcacggc gccgggtcg aggagcgtga cgcccgccgc    7620
cagggtcttg tcgcccagga cgaagcccag cgtgacgtaa tcggcggccg cccggaggct    7680
gagagccggt ttgcgcgccg aggtcctgag cacggaagcc agatccgatc cgaacgacag    7740
tccggccgcg tcgacccgcg agtagatcgg atacgagccg aacggatccg ttgcaagcag    7800
cactcgctct cgcgccggat cgatcacggc gagcgaaaac gcgccgtcga gccgggcggc    7860
catggcaggc cccagccgtt cgtacaccaa agccaggaga tcggcgcgcg tcccgttcgc    7920
cgaaccaaac tcgctccgga gcgccgcttc gttgtagaga tcgccgcgaa agatcaccag    7980
cggccggctg acgtcgcgcc cggccgccac ctcgtcgagg aagccgagcg tgccgagccc    8040
gaccgcccat cggtccgacg gatctcgata gacgaccgtg cggcgcggcc tgcccgactt    8100
cgccgctgac gcggcgtcgg cgcggctgaa ggcggtagat gccggctcat gccggcaaac    8160
gacgccgagg aggccgttca tcagacgacc ccgcgggcag acacgcgcga cggaacaaga    8220
ggtgtcgtct gagcggccgg aaccggccgc gcggcgtgcg ccgccaccat ccggttgagc    8280
gcaatccaca agccgaggaa gtgccacagg atctcggtgt actgaaagat gacgaacgtt    8340
ccgcccacgc agaatgcaat gagggcggtc tggacggcca gcgcgtacgg cgccagctcg    8400
gccagggccg gatcccgctt cgagacccgg cggaccctgc ggcaggcgat gagcgccagg    8460
aggagagtga gcaggaacgc cgcgaggccc acgaacccaa gctcagccac gagacccgaac    8520
cacgcgctgt gaaccgagcg ccccttgccc caggcgccgt ccgtaaatcg gccggtcatc    8580
ggccatgcgg    8580
taggcggccg tgaaggcatt gtgaccgacg ccgtaaatcg gccggtcatc ggccatgcgg    8640
agcgcgacc tccagaagtg cagccggctc tgggccgatc gctgatcgcc aaccgccgtg    8700
tcgctcgcgt cgatgctcga gtcggacgag ctgatcgagc tcatcgatc ccagaactct    8760
tccgtcatca cgagcgacat cgcgacgaag accgccgccc cggccacgag gacgagcatc    8820
ttgcgcttcg agtacaggaa atagacgagg ccaagcacg cgagcgtgag gaagccgccc    8880
cgggagtacg tgctgatgcc gcgatagagc acccgatga gcagcacgcc cagcgcccac    8940
ttggcccatt tcctcgcttc cgtctgcatc agcgtgacga tgagcggcac gagcatgaac    9000
```

```
atgccgacgg  cgacgccgtt  gttgtcaccg  agcgagtaga  cgtcgttgaa  gttcttgatg    9060
ccgggaatga  acaggagctg  accccagccc  tgcttggcgg  cctcgaaccc  gagggagagc    9120
gagatcacga  ggaagacgag  ccgcaaccgc  ttgacgtccg  tcgtgaggac  tgcgagcagg    9180
tacgtcatca  cggtggactt  catgaagtcg  atcatgtagc  cccacgcgaa  gtcctggtag    9240
ggcgacgcca  gcgtcgacga  caggctgtgg  agcccgaaga  aggcgaggag  cgccagccgc    9300
gcgtcgatcc  gcagcgcctg  cccggacagg  aacgcgatcg  cgagcgtgta  catgcccgcc    9360
aggagcgaga  tgttcatgtg  gatgagaaag  tcgctccaca  cccagagctc  cggccggaag    9420
tacgcgatga  acaggtagaa  cagaatcgcg  tagaacggcc  cgcgcagcgc  gtggaaggcc    9480
ccgaaggcga  ggagggcgag  gacgaaagcg  gctcgaagca  ttgccgtgtg  ttacctgtta    9540
tcggtcccga  tggtccagga  cggctgcgtc  cgggtctcct  ctacattgaa  caccttgtag    9600
agggccggaa  tcgaggtgaa  catgagcacc  acgaaaagaa  tcgccgacac  caccatgtag    9660
gcgaagaacc  cccgctgctt  gtagagcttc  tccggattct  gcacggggct  gttcggcagc    9720
atcccgatgt  gcaggtagta  ggcgaacatg  cccgccgcca  ccggcgcgaa  caggatcagc    9780
tcgaggtgat  agcggacgat  gaagatgccg  ccgaagaggg  cgcagcaggt  ggcgtagaac    9840
accatgctca  cgagcagccg  ctcctcgttg  tagaaggcaa  acgacttccg  gtacgaggcc    9900
gcgaccgacg  cgttgccgat  gtgccggaac  tccgcgtacc  gcttcgtcgc  catgaagaac    9960
gcgcccacca  tccagtacga  gatcacgagc  gacatcggcg  gcacgcgatc  ctgaatcagc   10020
gggaaccagc  cgaggaggag  ccggacggcg  ttgttcgccg  actcgctcag  cacgtcgagg   10080
tacggccact  ccttcgtgcg  aatgggcggc  acgttgtacg  tgacgccgag  cacccagagc   10140
gccagggccg  acagcgcgaa  gtatctgttc  acgagcagcg  cgagcacgaa  tccggccacg   10200
cccacgagga  tccactcggt  gtagcccgcc  gccggcttga  tcttcccgga  gggcaccggc   10260
cggtgccgtt  tctccggatg  cagcagatct  ctcggccgt  ccaggagctc  gttcagcacg   10320
tagttgctcg  aggcgatgag  gcaggtcgcg  gcgagccgcga  gcgcgaggcgg  cgggacggcg   10380
gtccagccga  agagctgcgg  ttcgtagaag  aacgccagca  gcacgccgag  cagcatgaac   10440
gcgttcttga  accagtgatc  gatccgcgcg  atctggacgt  acggccaaat  ccgcgaaccg   10500
gagctagcgg  ctgacgacat  cgttgagctc  ctcgagccgc  cggatcgcga  accatccgcg   10560
gccggctgtg  acgccgtggc  ggccggcgat  gagggcgcag  gagagtcctg  cggccatcgc   10620
tccttctccg  tccacgtccg  gtcgatcccc  gacatacagg  acctcccgag  gtccgagtga   10680
ccacatctca  cacgccacat  ggaagccgcg  cgggtgcggc  ttcagggcgt  tcacgttgct   10740
ggcggtcgtg  cagagcgcca  gcgaaaaatg  ctgcgacgtg  ccgagcgcgg  cgagcttgct   10800
ctcgtgcgcg  tagtccgaca  gcacgccgag  cctgacccc  gccgcgcgaa  gcgagtcgag   10860
cgcgccgagc  aaccccgggc  gccggcacca  ccgcagatac  ttgagcggcc  ggcgcaccat   10920
ccattcgttc  accgtcgccg  ccacgacctc  gttgtccagc  ccgagcgcct  gcgcggtgcg   10980
cgcgatctgg  cgcgccgcga  gcggctcgtc  ggccgcgccc  aggcgccgga  gatcctcatg   11040
cgcccggcgg  tactcccgca  cgatgcgcgc  cgtgtccact  ccgcgccgcc  accggaacgt   11100
cacgagcggc  gtggcggcga  gctccgctgc  catggcggcg  cgcaacggac  cctgccggta   11160
cagcgtgccg  tccacgtcga  acagcacggc  cttgacccccg  cgcaacggct  tcactctgcg   11220
aggcccagca  tccgcagaat  cgccgacacc  tttgacgcgt  cgcgaaggaa  cacgtcacga   11280
acccacacgg  cagccagccc  cgccaacagc  accgccgctg  ccgcgcccgc  tgccgtccgc   11340
cagctcccgc  gccaggcgcg  cgtgtcgatc  gcggcggcgc  cgtagcagag  cagaatcggg   11400
acgagcggca  ggtgataccg  gggatgtccg  aagacgatcg  cgtgaatgcc  gctcatgaac   11460
acgatcaggc  tcaggagcag  caggtgcgcg  cgccagttga  ccggccgcgc  gaacgagatg   11520
ccgatgacgc  cgagcaccat  gaccgccgcg  tacgatccca  tgatggacac  cgcccggcg   11580
aagaacagcc  accgccgcgg  acggtagaac  ccatttccga  cgccggcgat  gaagtcgccc   11640
tccagccccc  agaagtctgc  gaacttcagc  accgaacggc  gaagcgtcgt  tcccggatgg   11700
gccttcatgt  agtcgagcgc  ctggcgctgc  gcccacttct  ctttcgtgcc  ctccgtccag   11760
cggctgccgt  ccggcgcgcg  aggcggcatc  gtgcgcgacc  attccttgtc  gcccgtgagc   11820
gagatcgcgt  cccacatgcg  atcctcgggc  gtgtacgcgt  agttgcccat  catcaggttc   11880
agcccgccga  gcgtgtccac  gaccgtgaac  gagcgttgca  ggagcgtgtt  gcgcacgctc   11940
cacgggccga  cgatcgcggc  gtatccggcg  aagagaagca  cggccacgcg  cagacgcgcg   12000
cgaagcggga  tgcccagacc  gacgagcgcc  gcgccatga  ggatcggcac  gtacggccac   12060
atgatcgagc  gaacgacgcg  ggacgcgcgccg  agcgcgacgc  ccgtcgcgag  cgcgatccac   12120
gccgagtgcc  gccctgatgg  gcgatcgatg  agcgtcagac  atcccaagag  caacaaaaga   12180
aggagtgtga  tgaagagcgt  ctcggacagc  accagcacgc  cggagaacag  cagcgaggga   12240
tagaacgcga  acgccgccgc  ggcaatgagg  ccggcgcgtg  cgctgaacaa  cctgcgcccg   12300
atgagataca  cgagccagac  accgagcagg  ctcaacggaa  cctgcgccag  ccggaccgcc   12360
gtcagcgatt  ccgagcccgt  gaccgcccag  acgcccgcga  cgaacgcggg  aaagagcggc   12420
gcccggatgg  acgtcggttc  ccccggcccc  cacgcgaacc  cgttcccgc  gacgacgttt   12480
ctggcaagcg  tcgcgtagtg  ctgctccgtc  cgaatctcga  gcgcgacgtc  cttgaacccc   12540
caaatcaggc  cgagccggag  gaccaggcc  agcgtgagga  tgacgagaat  tgccccggcgc   12600
tcccaatggg  tgcgctcggc  ctcgtcgcgc  ccggggggcga  cgattggcca  ggggggccgac   12660
gatgggagtg  ggggagcctg  catcattcgc  taaaacaagc  agttacggtc  aaacctgtgc   12720
caccggggaa  tcagcaaaga  ctctgccaac  gggtccgccg  gcggcgacgc  ccggcgcgcg   12780
ttttccggcc  ggaagtcggc  ccccaaggcg  accggctgg  gcacagccca  gtagaaatct   12840
cattctggga  tgactttttgt  ccatgatcca  tgaccatttc  gaggggcctt  cgtgacgccc   12900
tggggatggc  ataagcttgg  caccgccgat  cgacgtgt  cgatgatcag  cggccgtgag   12960
gagctccaga  aggcctatcg  ggacgaccgg  gtcgccgtg  aatacgtggc  gcggcggttc   13020
cagtcaccgc  tggggggcgct  cctccattcg  cgtcagatcg  gggtcgtgcg  agagctcgtc   13080
cgcgcgcagg  gcatccggcg  cgcggcggag  atcgcacccg  gtccagcgcg  gctgaccgtg   13140
```

```
gacatcgcgc cgctcctcga ccgcgtgacg ctcgtcgacg cgagcgccca gatgcttcac   13200
gaagcgcgcg agcggctgcg cgaccggacg ctcgccgcgc gggccaatct cgtgcaggcc   13260
gacgcgtttc gtctgccgct cagcgggccg ttcgatctcg tctacacctt caggctcgtg   13320
cgccatttcg agcgcgccga tcggctgcga ctctaccgcc agattgcgac gattctcgga   13380
cccggcggct ggctcgtctt cgatgcggtg aatgagattg tgtctgcgcc gcttcgcgcc   13440
aacgcgaagc ccggcgagta cgagcactac gatgcgctgc tcacaccggc cgggctacgc   13500
gacgagctgg ccgaggccgg gttcacgatg gtgtcgctca ccggtgcgca tcgccggttc   13560
cccacgctca tgaagtgcca gatgtatctc gcgccgaggt caacggtgct cgcgcgcgcg   13620
gcgatggagg tgatcgatcg ctcggcggc gaaccgctcg agtggatcgt cgtatgccgc   13680
cgcgggtgac ctactggacg ggtacctggg atccgatcaa agaagcgatt tcgaaggaag   13740
tgaacgccct ccgcaccggg tcgcgggcga acgcgacggt cgtgtcgttc tcgccggccc   13800
agcgcagccg tctcgtgccg aaagatcgcg tgctcgtgtt gtccgtcgc gcctggatcg   13860
cccttcgggc cgcggccgcc gtgctcgagc cgcgcggcga cgtgacgcac gtcttcggcg   13920
gcgggttctc ctggcatctc ctgagggcgc tggggcggcg gccggttctg ctcaccgccg   13980
tcaaggaaga tggcgctgcc gatctgcccg gggcgatcgc gcacgtggcc gtcgaagtcg   14040
gccgcgctcg cgacacgtgg atcgcggccg gtatgtctcc cggccgcgtg cacctcattc   14100
atcccggcgt ggacctcgac tggtatcggc cggtgccggt ctcgtcggag agattcacgc   14160
tgctcttcgc cagcacgccg tccgatccgg ctgagatcga gccgcgcggc attccgctcg   14220
tcgtcgagct ggcgcgcggctc cgtccggaga tcgacgtgct ggtgccgtgg cgccggtggg   14280
gcgacgtcga cgcctccagg cgcgcgctcg acgccctccg tccgccaccg aacttcatcg   14340
tctcgtacga ggacgcaccc gacatgcgcg cctgttacgc gcgggcgcac gccaccgtcg   14400
tgtgcttcgc gggcggcacc ggcaaggcct gtccgaattt cgtgctcgag ggattggcca   14460
tgggacggcc gtgtctctcg acgcgcgaga acggcatcgc ggagatcctc gatcgatcgg   14520
gcgccggcat catcgccgcg cgcgatgccg cgtcgctcgc cgcgagcctg gacacgctcc   14580
gcgccggctg ggccggctgc gccgagcgcg cgcgccggct cgcagaggag acgttcgacc   14640
tcggccgctt ccgcgccgcg tacgacgcgc tgtacgtgga ggtggccatc gccgccgctg   14700
gcgagcacgc gaggcgacgc tcttgacgaa gggactcgcc ggacggctgc tcgggtgggc   14760
tcagaaggcg aacgattcg ccgactacgt ccgccatcga tcgcagatga agcgcgacgc   14820
gcgggcgaat cgccgcaggc tccatcgggc gctcggcggc cgcgtgaacc ggatcctcat   14880
cgtgcggctg ccagcatcg gcgacgtcgc gcgcgcgacg ggcgtcctcg gggccctgcg   14940
cgcgcagtac ccgcaggcga cgatcgactt cctgacgtcg gacgcgaccg tggccgtcat   15000
ccggcaccac cccgcgatca acaccgtcta cacgctccag gatctgtccg gcctgtccga   15060
gtacgactgg atcatcaacc tccagaacct cgtgccgccg gcgagctttc tcagaacgtc   15120
cgggctctcg tacccgcaga ttctcgagca gctctcgagc cgcagccgtt ccaggatcat   15180
caccggccgt cacctcgaga acggctcgga gtccacgccg acgaacatcc tctactgcat   15240
ctgtgagatc gaggagcatt tcctcacggc gctcgtcccg tacgacacgg cccgctaccc   15300
ggtgaccgcg attcacgtgg acgcacaggc acgagcggcg gcacgatcga agttcgcgtt   15360
cccggagaac cgccaggtgc tggcgatctt cctcgggtcc aactcggtcg ggtgcggcgc   15420
cgatgagggc tttcgcacgt actccatcga ctacgtgaag cgcctccgtcg gtcacttcga   15480
agatcggttt gcgatcgcca tcatcgggca gtcccaggtg cgcaacccgg ccgaacgcgc   15540
acagtaccgc gagatgctcc gccgccatcc ggacgtgatc gatctcgtgg acaagacgtc   15600
gctcgacgag ctcgtggccc tcatggactc gtttttccctg ctcatcagtt gcgactcgag   15660
ccccgtccat ctggcgatgg cccgcgggagt gccggtcatc gggctgtacg tcaacgacgg   15720
gacgttcagg atgcacccga cgctcggcga cgagcgcttc gtggccatca acagcacgcc   15780
gccgtgcttc acctactcgt ggcgatggaa gttcttctgc tcgacgtgcc gcgacccggc   15840
gaccagatcg cattactgtc acaacgaggc gttcgtgttt ggcgtggacc gcattcggct   15900
cggccgcaatc gaccgcgcgg cggcgcggtt gctcaagggg agctccgatg caggacctca   15960
ttcgcgcgca gtttgaagag tcgatcgaga cgaagcgcaa agcgctctcg ctcgtcgatg   16020
cggtggccga gtcggcgcgc ctcctcatcg aggcgctcga ggccggacac aagatcctcg   16080
tgtgcggcaa cggcggctcg gccgccgact cgcagcattt cgccgccgag ctcgtgggcc   16140
gcttcgagaa ggagcgccgc gcgttgcctg ccgtggcgct cacgaccgac acgtcgatcc   16200
tcaccgcgtg gccgaacgac tacgcgtacg agaccgtgtc agccggcagg tgagtgcctc   16260
ggagcgccgg gcgacgtgct catcggcatc acgacgtccg gcaactcgcc caacatcctc   16320
gccgccatcg acgaggcgga aaagaaacgc atgaaacgaa tcgcgctcac cggacgcggc   16380
ggcggcaaga tgaagagtct cgccggcgtc tgccaggtca tcgtcccgtc agatcagacg   16440
tcgaggatcc aggaagtaca cattaccgtc atccacatct gggcgaagtt gatcgacaaa   16500
gcgctcagct gagctggcgg tagacgtcga acgtgcgatc ttcgaacgtc gaccaggagt   16560
gctcgcgggc gatgaggtcg gcggcggcgt cggtgaggcg cggcggagg tgctgatcct   16620
cgaacaacag ttgcagtcga ttcgcgtagc cctcggcgtc gtccgcctgg acgagaaagc   16680
cggtcttccc gtcttccacg acttcacgca cgcccgggat ggtggagccg atcacgggca   16740
cgcgggccga ctgggcttcg agcaccggga gcgcgcagccc ctcgtgcgtt gacggcagca   16800
ggaaggcgtc ggcggcctgg agcagcaccg gcgcgtcctg gcggaagccg agaaagcgca   16860
cgcgatcctc gagcgcgagc tcggcagaga ggcgcctcag cgcctgctcg tgttccttgc   16920
cctcggtccg gtcgctcccg accagccagc actcgaccgg caggccgcgc gacatcagca   16980
gcctggtcgc ccggattgcc gtagcctgcc ctttgtgtgg cgcaagattc gcgagcatga   17040
ggacgacgaa gcggtcggtc tcgaggccgg cggccgtacg cgcggcttgc cggtcgccgg   17100
ccacgaactt ctcccgatcg accgcgttcg gaatcgcggt gaccggcacc ttgtcccggc   17160
gtccctgctg cgccgcgtcg atcatcacgg cgatgtagcg cgcacacgtg acgatgtgct   17220
gcggcgggct cttcaggacc caggcgatct cctcgggatc cggttcgatg tgaaagtgca   17280
```

```
cggccacgcg cgcgcgcgcg gcctgcagcg cgggtctcat gaatccgtac acgacaggat    17340
tgtgcacgtg gaccagggc cgcccgacga gacccaatcc cggaagcatg cgcgcacacg     17400
cggcgagctg ttgcaggccg ccggcccgca gcgcgttcag gtcatacatc cgccgccgga    17460
cgccctccgc gtcgagcgcg gcgccgcgca gaccctcgc tgtaatccat gccgtgcacg     17520
ggacgccgcg gcgtttggcg gcggccgaca accggatcgc cagcaccgcc gcgccgccga    17580
ccagatgtga gacgaggatc tcgtgcaggc ggggggactg attcatggtg gcgcgcgatg    17640
cggccgggcc gccagcctac gttactcgcg tcgcaggtcc gcccgcgcgg tagagcgagc    17700
gcagcaggct gcgaagtccc gacgggagga agcacacggc gtaccgcgtg agatcggtgc    17760
gccgcatcgg caggcccgcc tgaccgaacc gtttccgcgc cgcgtgaaac tcgccgcggt    17820
cgtactcgag gtggccgagc gcgagcagga gacgatggcg atgatcgctg gcggcgcgat    17880
gcgccgactg atcgctcccc cgcgccagga accggtccac cgtcctgacc atgccttcga    17940
gcatcacgat ggacgacttc gacatgttgg tgtcgtgcag cctgtaatgc gtcaggggtt    18000
ccgggatgaa accgaccggg cccttcgccg agaggcgcag ccagaggtcc cagtcctcga    18060
cgccgtggat gccggtcgcg aacatctcct gccggagcac ggcggctttg accatgaccg    18120
acgaagcggt cacggcgttt cgcaggacga ggtcggggag acagcgcccc tcgagcggcg    18180
gaagatcgag cgtgcgcagg acctcgccgg cctcgtcgat gacggcgcgg cccgtatgaa    18240
tgaagctcca gtcggcgtgc tgtcggaaga gatcgacctg ccggcgaatc ttggccggat    18300
cccacgagtc gtcggcgtcg agcagcgcca cgaattcccc tcgccgccgc gcgatggcgc    18360
gattgcgggc gccggcaagg ccccggttgg cctgagagat caccgtgatg cgcgccgccga    18420
atctcgccag gatgtctgcc gtgccgtcgg tcgatccgtc gtcgacgacg atcacttcga    18480
cgggctgata ggtctgggag agcgcggact gcacggccct ggcgatgtag cgctcgccat    18540
tgaacaccgg gacgatcacc gacacgaggc cggggatatc ggtctgggcg gccggcgagg    18600
gctctcgagc ggcttcagcg gccatggccc tctgcactgg cgtcgcgagc gaacagcatg    18660
acgtttctcg aaagccacgg agacaacggc gcgctcttca tccgctcccg cagcgtctgc    18720
gatgccgggc cgaggtagct gaacccggcgt gccgcgaaga tcttctgcca gtactcgagc    18780
ggctgttcgt tgatgtggcc cattccgccc tggccccggct gcgcggcgct gaagagcacc    18840
agcggcgccg tgttcgtgag cacggcgacg agcttctcgc caagcggcgg cggaaggtgt    18900
tcggcgacct cgaggctcat ggagaggtcc gttgcgtccg cggccgtgga caccgccgtg    18960
agatcgaatg agttcagcgc gacgccgagc ttttccgcgg cgtaccgtcg cgcgtactcg    19020
gatcgctcgt agccggccgc cttcacgcct cgtgcgagga acgcgtgcac gaagtgaccg    19080
gtgccgcagc cgaagtcgtt catcgatcgg atgccgggtc gcaacgccag gaccccatca    19140
acgagcacgg gcgccgcctc ggcggcgacg tgcccgaaca gatccatgac ggccgcgttg    19200
tacgtccatc gctcgaagcc cagccgttcg ccgatccggg acatcttcag tccggcctgg    19260
caacgcccg accagtagag ccgcttgagc tgggtcacga tcatggggttc gtcccgggcg     19320
cgacgcgaac gcgcggcggc gcatgcagcg acgccgagaa ctccacatcg gccacgccgg    19380
cccacgatcg atgttccttg accgccagcg tcgcaagctg ccgcgacggc acgaggaacg    19440
tctgggtcgg gttgtgaaac acgtgcaccg acaggtagta gtggccgcga acgagatgcg    19500
cgcggaaatg gaggtcgagc gagaactcgc ccatgaacgc ggacgcgtcc ccgagatccc    19560
gccggtggat gttgccgtcg tagacgacga gctgatcggt ggacgtgtac accatgaagc    19620
cgaacgtgac gtcgtccagc ggctcggtac ttctcgcggt gacctggagg ttcagctgct    19680
tccccggcgt cacttcgctg tgcgtgcccg acgactcgtc gccgacgaag agcgagccgg    19740
acacgatctc gacggggctc ccgacactcg gctgttccga tttcctggac gcgcggacgt    19800
acgcgtcgag gacttcgatg gccgtgccgg cggccttcac cgagccggcc agatacaccg    19860
catgctcgca gagggtcgcg accgcctgca ggttgtgcga gacgaggacg atcgcgatgc    19920
cctggcgctt gaactcctgc atgcgatcga agcacttctg ctggaacgtc atgtcgccga    19980
cgctcaacac ctcgtcgatg atcagcacgt cgggccgcag gtgcgcgagg atcgagaagc    20040
cgagccgcgc gttcatgccc gacgagtagc gcttgacggg cgtgtccacg aagtcggtga    20100
cgccggcaaa ctcgacgatc tcgtcgaacc gctcggcgat ctcgacccgt ttcatcccca    20160
tgatcgcacc ctggaggaac acgttctcgc gcccggtgag gtcgggatgg aaccccgcgg    20220
cgacttcgat gagcgcgccg gcgcgccccc ggacgatcga gcggccggtc gtcggcttca    20280
ggatcttcgt cagcaccttg agggtggtcg acttgcccgc cccgtttgcg ccgatgatgc    20340
cgagcgcctc gccgcggtgg acttcgaacg agacgttct gaccgcccag aactcctgtt     20400
cgtccagcgc gtcggcggga cgcggttgaa acagcccgcg cacgagccgc ggcacgaggt    20460
cccgcaggct gtcatgacgc tcgcccccgc ggaacttctt ggagacgttg tcgaagacga    20520
cggggccgt cacgtcagac gttctccgcg aattcgaact cggagcgatg gaagaacagc      20580
cagccgaaga cgagcgtcac gaccgagacg atcgccgtga cgcccagcat cacggacggc    20640
ggctggttca ggagcagcac ggcgcggaag ccgtcgatga tcggcgtcat cgggttgagc    20700
tgcagggccc actgagctt gccgcccacc atttcgaccg ggtagaccgt cgacgacgcg      20760
aacatccaga cggcgatgac gacctcgaag aggtatttca cgtcgcggta gaagaggttc    20820
gccatcgaca ggatcagggc gatggccgtg gtgaatatca cgagcacggc gagcacggcc    20880
ggcacgcaga ggatgttcgg tccgaccggc acgtggtagt agatcatgag cgcgacgagc    20940
accacggagc cgatcgcgaa gtccacggcg ctcacgatca cggcggagaa cgggaagatc    21000
tccctgggaa agtacacctt cgagacgaga ttcgtgttcc cggtcaggga ggtcacggag    21060
aagcggagcg ccgaggagaa gaagttccag acccacaggc cgcagaacgc gaacaccgga    21120
tacgcgaccg gcgtgtcgat ggccgcgacg cgcatgaaga tcacggagaa gacagccgtg    21180
ttgatgagcg gcatgaacac ggcccacccg aagcccatca ccgactgctt gtagcgcagc    21240
agcaggtcgc gccgcgtcat ctgcatcagc agctcgcgat actcgaactg ctcggcgagc    21300
atctcacggc aatcggccag gaaggcgcgc atcgtcagcg gtgtcagtcg cggcgtcgtg    21360
ccgggcggat catgagggca ggtagcgctc gagctgcccg aagagcaact gcacgaactc    21420
```

```
ggtcgcgcgc cgctcggcgg cgacgtcgcc gttttcgctc gacggcacgg ggctcgtcac      21480
gcggacgagc gcggcgtccg atcggttcag ccgcatcgcg tcgtacacca tgtagatctt      21540
gctccagtac tcgctggcga tcacgcggcc gtggctctgg taccagtaca gcacgatttg      21600
ccgctcgatc cccttctgga tgacgtatcg gttgatctcg atgggctcgg tgcggcccgg      21660
caccttgacc agcgcgcggc cggtggatac gggctcccag ccggcgcccg gcaggcagtt      21720
catcggcgag tggatcgtgt cgccttcgcg ctggctctgg tagtagccga tgtagagcga      21780
cacgtagggc tcgcgcggcg ccaggtagat ccggttgatg tactcgtcca cgccgagcac      21840
cgccatgacg tccttggtga agggcgcgct gtcgcgtccg gtccaccgct ccatctggaa      21900
cgggacgttc gcgagcggct ggcgaagcgg gacctgctcg gcgcgcatcg agcgcgcggc      21960
aaacgccgcc gtgccgagga aacacgct gagaatgatc agtcgtctgg tcatgaatct      22020
ggactcacac cgtggccggc gcggtcaccg ctcggggcgg aggcgtgaag cgccgcagga      22080
cccagaccac gacgaacaac atgacgaacg cggcaagaaa gacgagccag cccgagaacg      22140
tgtggaagaa cccctgcgcc gcggcctcgc cgtagtagtg cgccgccacg cccgtgccgg      22200
ccacacgcgc gccgttcgcg ataatcgcga tcgggatcgt cgagacgcga atcgccaggc      22260
gcatgccgac gcgcggatcc gtgaagtagc cgtacacgat gcccagcgtg agcagcgaga      22320
tgagcgaccg gatcccgctg cacgcctcgg ccacttcgag cgtcgtgtgc gcgagcacga      22380
tcacgttgcc ttcgcgcagc accgggatgc tgagcgcgga cagcacgccc tcgcccacgc      22440
gagacgccag gagctgcagc ggaaatgcga tctgattgaa gatgatcgac gggatcggga      22500
tcatcagcac caggaacgcg agcggaaacg cgagcaccca gagcttgcgc cagccgagca      22560
ggaacacgat cgcgccggta agcgaggcta gtatcgacac gcgcgtcaga aagagctcgg      22620
cgccgagcaa ccccgcgacg agcagcgcga gcccggcgag cacgatgacg agaccgagcg      22680
cgctcggcgt gtccggcagc gccgagagac gcgcccggcg ctcccaggca aagtaggccg      22740
cgagcggtac gatcagaatc ccgtgggagt agttgtcgtc gttgatccag tcctgcgcca      22800
accgcacgat cacctgccca taaaggagca ggaaactcac ggcggcgata ccgagcggaa      22860
ccagtctggc ctgaagcaca ctcatagcga tcgggacgaa tatagcaaag ggctctcccg      22920
ggcagaactt gggcgccgaa acacgtaacc atcacgctca caggaagtta gcggctaaat      22980
gacgggattg gccggctgga ggggcctcgt gaatctaccc aatggatgac aaatatccct      23040
atcaactcct cctttttgact gcgtagcacg gaagatcatc ggcctcttca gatgcaggac      23100
aaatgtcatc tgtcaggaag agtaatgtct ctggccacag cgccggtcaa gtgaaatccc      23160
ccgacagaat cgcgtaccac cgcgactggc atccggcgag aaaagtcgcc tcgtcggccc      23220
gaaaactgca ttacaggaac gccaggtcgg cagccggttc gttttccgtc agtatttttg      23280
ggcattgcag gcgctttcgt ctgccggggg gcggaaccgt gcgaccgggc cgatcacgcg      23340
tggagcaggg cagggacagg catggatggt tccgtgagac agccgcgttc gggagacggc      23400
cgggctcggc gggagctcgt ccacctggcc gtggaccgcg cggcgcgcca atttcctgac      23460
gcgccggccg taagtttttga cggcgtctcg ctctcgtatc gcggcctgac tcgccgcgcc      23520
aacgcgctcg cgcaccgcct gcagcgtcat ggcgtcgatc ctgacgtggt ggtcggcatc      23580
tccctcgacc ggtccatcga gctcatcgtc gggatcctcg ccgtgctcaa ggccggcggc      23640
gcctacgcgc cgctcgatcc tgcgtatccg aaggaccgcc tcgagtacat ggtcagggac      23700
gccggggcgc ggatcgtgct cacctcgcga cgcgtgagcg gtctgatccg cgtcgatggg      23760
gtcgagctca tctgcctcga tgcggacgag cccgggggcgt tggcggaaga gggccatgcg      23820
ccggacgccg gcgtcgatct cgaccacctc gcgtacgtca tctacacctc gggatcgacc      23880
gggcgcccga agggcgtcgc gatgccgcat cgtccgttgg cgaatctcat cggctggcag      23940
gtcgacgcat cggccgtcgg cgccggccgcg cgcacgctcc agttcacttc gccgagcttc      24000
gacgtctcgt tccaggaaat ctttgcgacg ctcgcgaccg gcggcgagct cgtgctcgtg      24060
tccgaggaca cgcgccggga tccgcgcgcg ctgctgcggg tgctgcgcga tcgatccgtc      24120
gcccggctgt atctcccgtt cgtcgcgttg cagcagcttg ccgtcaccgg caccgacgat      24180
ccggatctgc cggcgttgcg tgaggtcatc acggccggcg agcagctccg ctcgacgccc      24240
gcgctcgtcc agtggttcgc gcggcatccg cagtgcacgc tccacaacca ctacggcccg      24300
gccgaagcgc atgtcgtcac tgcgcacacg ctgactagcc ccccgcggga atggccggcc      24360
ctgccgccga ttggcgtgca gctgcccggc gtcgagattc tgatcgtcga cgagcatcgc      24420
cagcccgtcg cgcacgacga cgtgggcgag atcctcattg gcggcgtgtg cctcgctcgc      24480
ggctacttga atcagccgca gcgcacggcc gaacggtttc tctcgcatcc gcttcgcccc      24540
gatgctcgcg tgtatcgcac cggcgatctc ggccgtgtgc gcgacgatgg cgcgatcgag      24600
ttcgtcgggc ggacggacga ccaggtgaag attcgaggct tccgtgtgga gccgggcgaa      24660
atcgagaccg tcctccttca gcatcccggt gttcgggaga ccgccgtcgt cgcgcacgac      24720
gacgcgtcgg gcagcaagag gctggtcgcg tatgtgattc cgtcagccga tcaggcggcc      24780
accggagccg aggcttcggc acaggtctcc ggatggcaga ccgtcgtggga ggacacctac      24840
cgccggccgg cggactcgac cgatccggcg cttgccgcga ccggctggcg gagcagctac      24900
acggggatgc cctacaccgc ggcggagatg cgggagtggg ccgatgagac cgtcgcgcgc      24960
atccggtcgc atgcgccacg cgaggtgctc gagatcggct gcggcacggg catggtgctc      25020
ttccgtctcg cgcccactg cgctgtctat agcgcgacgg atgtctctcg cgccgcgctc      25080
gaccacgtcc gggcccacgc gacggcctgc ggcgccagcc acgtcgagct cctctcgcgc      25140
gaggcggcgg actttaccgg cgttccggcg gccgcgttcg atgccgttgt gatgaactcg      25200
gtcgcgcagc acttgcctgg cgtggactat ctgacgcgcg tcctcgccgg cgccgcccgc      25260
accgtgcgtc cggccggatt catcttcgtc ggcgacgtgc cgaacctgcg cctgctcgag      25320
atgtttcacg cgtcggtcga gctcacgaaa cgcgccggcg gg atcttccgat tgccctcgtg      25380
cgcgagcgcg tgcagcgtca gatggcgctc gagcgcgagc tcgtgatcga tccgatgttc      25440
ttcgacgccg tcatggaccg catcccggc atctcgggag cccaggtcca gatccggcag      25500
ggtgccttcg acacggagat gaaccggttc aggtacgacg tcgtgctcac ggtgggcgga      25560
```

```
cggccggccg ggcgccgtcc gtccgtcgcg ctcgactggc gtacggagcc ggtcgccgag   25620
cgtcagctcg ggcaactgct cagcgccgat gcgccggacg cgatcgagat caggaacgtg   25680
accaacgcgc ggctgcgtca gccggccgcg tgcttcgatc tcgtcaagcg tgccgactgg   25740
accggcacgg ccggcgagac gcttgagagg gccgccgcgg acacgcgcga cgcgatcgac   25800
ccggaatcgt ggcggcgtct cgcgaacgcc gccggctacg acgccgagtt cacctggtcg   25860
gccgagatcg gcgccatcgg cttcgatgcg cggttcgagc gccggggcac cgcgtcgctc   25920
gccgctcccg tgactcgtca cgcggttccc gccgcggagc cgcggggcat tcagcagtac   25980
gcgacgaatc cgctcccgca tgcggtggcg cggcgcctca cgcccgagct tcggcgattc   26040
ctgcaggagc gcgtgccgga tcacctggtg ccgtcggcct tcgtcgtact cgatgcgatg   26100
ccgctcacgc caagcgggaa gatcgatcgg cgcgcccttc cggcgcccga gagccgccgt   26160
cccgatctgg acgtggcgtt tgccaagccc gccacggagc tcgaacgcaa gatcgcgcag   26220
gtctggcaga cgacgctcca gatctcgtcg gtgggattgc acgacaactt cttcgatctg   26280
ggcgggcact cgctcctgct cgcgcaggcg ttcgagcgta ttcgctcgct cgtgccgggc   26340
aagcagtggt cgatgatcga gatgttccag tacccgacgc tctattcgct cgccgggttt   26400
ctcagtgagc ggccggacgc cgcggccgag cagcttgccg gcgcgcagga tcgcggacgc   26460
cgccagcgcg agatgctgac gcgtcagggg cctgccctga gcagagtcga aggggtccgg   26520
taatgcagaa cggcatcgcc ctcgtcggga tggccggccg ctttccgggc gcgcggtcga   26580
tcgaggagta ctggcgcaac ctcgtcgccg gcacggagtc gatcacggtc ttcaccaacg   26640
aggagctggc gacggccggg atcagccaga gcgatctgca gaatcccgac tacgtccgcg   26700
cgcgcgggct cgtcccggat ccggacctct tcgacgccgc gttcttcggc atcagccagc   26760
gcgaagccga gctgatggat ccgcagcacc ggctcttcct cgaggcgtgc tggctcgcgc   26820
tcgaacacgg cggcgatccg cgcgggtcgt tccccggtct catcggcgtg tggggcggca   26880
tgagcacggg gatgacgaac agcacgtacc tgctgtcgaa tctccacagc cacgcgggat   26940
tgctccagcc ggaggacgtg ctgccggcgc tgctcggcaa cgagaacgac tacctcacga   27000
gccgcgtctc gtacaagctg cacctgcgcg gaccgagcat gaacgtgcag accgcgtgct   27060
cgacgtcgct cgtcgcggtg gtgcaggcgt gtcagtcgct gctcacgtgg cagtgtgatg   27120
cggcgctcgc cggcggcgtg tccgtgtcgt tcccgcaaca ggaaggctac gtctacgtcg   27180
aggcggcat cgggtcgcca gacgggcgact gccgcgcgtt cgatgcgaac gcccaaggca   27240
cggtcttcag caacggcgtc ggcatcgtcg tcctgcgaag gctcgaggac gcgatcgccg   27300
acgggcagac gatctacgcg gtgatccgcg gatcggcgct gaacaacgac ggctcggcga   27360
aggtcagttt cgcagccccg agcgtggacg gccaggcgga ggtcgtggcc atggcgcaga   27420
gcgccgccga cgtggcgccg cagacgatcg actacgtcga ggcgcacggc accggcacgc   27480
cgctcggaga tccgatcgag gtggctgcgc tcacgaaggc ctttcgcgcc ggcggcgcca   27540
cggagaacgg gttctgcgga ctcggatcgg tcaagagcaa tttcgggcac ctcgactcgg   27600
cggccggcgt cgcgggggctg atcaaaacgg cgctcgcgct ccatcaccag cagattccgg   27660
cgacgctgca ctacgtcgag cccaacccga agtgtgattt cgccaacagt ccgttcttcg   27720
tgaacgccgc gctgcgacca tggccgcgcc gtggccatcc gcgacgcgcg gccgtgagct   27780
cgttcgggat cggcgggacc aacgcgcacg tcgtgctcga ggaagcgccg ccgccgcgg   27840
agcgaccggg cgccgagccc gccagcgacc tgctcgtgat ctcggcgaag acgccggcgg   27900
cactcgacgc ggcggctgcg aacctcgcga gtcatctcga ggcgcacccg gacgtcagca   27960
tcgccgacgt ggcgcacacg ctgcgcgcgcc gccgacagga gtttccgcat cgacgagcgc   28020
tcgcctgcac cagccgcacc gatgcgatcg ccgtgttgcg cggcgcggac gccaaacgtg   28080
tcgccaccgg cgacgccgaa tccggcagcc gtgccgtggc cttcatgttc tccggcggcg   28140
gcacgcagta cgcgaacatg ggccgggagt tgctgccggt cgcgcaggtc ttccgcgaag   28200
aaaccgagcg ctgcctcacg cacctgtcac tctccgccga cgtccgccgt gcgctgtttc   28260
cggatcatgc ggacgtcgag cgggccgcgc attggctcga agctccgtcc gttggtttgc   28320
ccgcgctctt tgtcgtgcag tacgcgctcg cccgtcagtg gatggcgtgg ggtgtgcagc   28380
ccgacgggct catcggtcac agcatgggcg agtacacggc cgcgtgcctt gcgggcgtga   28440
tgtcgctcga gaacgcgctc gcgctcgtgt gtctgagagg ccggctgttc gagacgcttc   28500
cggccgggga gatgctcggc gtcccgcttc ccgaggcaga ggtgcggccc gcgctcggcg   28560
atggactcgc gatcgccgcg gtgaatgcgc caagctcgtg cgtcgtttcc ggcgcgcccg   28620
acgcgctcgc ggcctttgcg cgcgtgatgg ccgaggccgg tgtcgaatgc cgccgcctcc   28680
acatcgcggt cgcggcgcac tcgccgatgg tcgatccgat cctggcggag ttcgagcgtt   28740
tcctccgcgg catcgcgctc gacgcgccaa gggttccagt cgtgtcgaac ctcagcggct   28800
cgtggctgac cgcgacgaa gcgaggtcgc cgcagtactg gacgcggcat ctgagacaaa   28860
ccgtcaggtt cttcgacggg ctcggccgac tgctcgccga gccgaatcgg gtgctcctcg   28920
aagtcgggcc cggacagacg ctgtccggat tcgcacggca gcatccacag aaaacggcgg   28980
cgcacgtcgt tcaagcgacg ctccggcacg ccaaggaacc ggcgccggac gtcttgttcc   29040
tgcagcaggc ggtgggcagg ctctggacgg caggggtgcc cgtgaactgg gctgcttaca   29100
acagaggcga gggcaggcgt cgtgtgccgc ttccggccta tccgttcgag aagaagcggt   29160
acctggtcga atcgccccag gcgccggtcc cagcgccgct tcaagccctg gcgccaaggg   29220
ccgatgctac tcagcagatg acgacaactg tgccttcgtc tccggcccga tcgcgtgccg   29280
atcgtatccg cgagcggctg acgaaatcc tgcacaaact gagcggccatc ccggcaggcg   29340
agatcgatcc cgcgctgtcg tttctggaga tgggattcga ctcgttgttc ctgacgcagg   29400
cgagtctgcg gttcaaggcg gagttcaagg tccggatcac gttccgccag ctcttcgaag   29460
aggcaccgac gatcgacgcg ctcacgagct acatcgacgc caagctgcct gcggacgcgt   29520
ttcctgcgcc ggcggaagcc gctccggcgc ctttcgtggc cgcaagcgtg ctgtccgcgg   29580
gccctcatgc ggccctggct ggcgtgccgg cggccgcgcc cggcacgctg cagcacgtga   29640
ttcaacagca attgcagttg atggccgagc aactccgcat gctgggcgga aatcccgccc   29700
```

```
tgctgcccgc gatgccgccg ccagcgtcgt ccgtcgcggc ggccgacgcg ggactgaagt   29760
cccgcgctac gacggcgccc gtctcgaagc cgcatgcggc cttcaagccg atcgacggca   29820
aagccagtgc ggagctgtcg cccgagcaga tggccgcgct cgccgcctgg atcgatcgat   29880
acacacggcg cacggccgga tcgaagcgac gcctggccga gtaccggccg gtgcttgccg   29940
atccacgctc ggtcgccgga ttcaaccggc tctggaagga gatgatctac ccgatcatca   30000
ccgaccggtc ggaaggatcg aaggtctggg acgtcgacgg caacgagtac gtggatctcg   30060
tgggcgggtt cggctcgctt ctgttcggac accgttcgcc gatcgtggtc gacgccgtgg   30120
aggaacaact ccaccggggg ttcgagctcg gaccgttgcc gccgctcgcc ggcgaggtcg   30180
cggcgctcgt caaggaattc acggggatgg agcgcgttgg attctccaac accggctccg   30240
aagccgtgct ggcggccacg cgtttgccc gaacggtcac cgggcgcgac aagatcgccg   30300
tcttcgaagg gtcgtatcac ggtctccttg acgaagtgct cagccggccg ctcgtggtga   30360
acggcgaatt gcggtcggcc cccgcggcg ccggcatcgc cggcagcgcg gtgagcgaag   30420
tcatcctgct cgagtacgcg aacccgaagt cgctcgaggt catccgcgcg cacgagtcgg   30480
agatcgccgc gatcctcgtc gagcccgtgc agagccggcg gctcgacctg cagccgaagg   30540
agttcctgca cgaggttcgc cgcgtggcgg acgagatcgg cgccgcgctg atcttcgacg   30600
aagtcatcac cggcttccgc gtgcatcccg gcggcgccca ggcgcacttc ggcgtccgcg   30660
cggatcttgc cgcctacggc aaggtcgtcg gcggcggcat gcccatcggc atcgtcgcgg   30720
gcaatgcgaa gttcatggac gcgttcgacg gcggccactg ggagtacggc gacgggtcgt   30780
tcccggaggt gggggtcacg tactttgccg gcacgttcat gcggcatccg ctcgcgctcg   30840
ccgctgcgaa ggccgtgctg acgtacatga gtcgcaggg tcccggcctg cagacgcggc   30900
tggcagatcg gatcgagcgg ctcgcgaacg acgcgcgcgc gatcgtggcc aggcacggcg   30960
cgccgtatca gattacgcag ttcagctcga tgatgagcct gaacttcccg cacgaccaga   31020
agacggcggt gctgctgttc ttcctgatgc gcgagcgcgg gattcatatc tgggagggcc   31080
gtcccttctt cttcacgacg gctcacaccg aggccgacta cgccgccatc ctccgcgcgc   31140
tcgacgagtc gctggccgag atgcaggcgg ctggcttctt cggcgcgccc gcccactccg   31200
cattgcaggt gaccgggacg gcggaccagg acgtcctgcc gttcaccgaa ggacagcagg   31260
agatctggct cggcggccag atgggcgagg cggcatcacg cgcctacaac gaagtcgtcg   31320
cgctcgacct gcgcgggcct ctgaatcgcg cggccgatgca gcgcgccgtc gacgaggtcg   31380
tcgcgcgtca cgaatcgctg cgcatgacgg tggccagcga accgctcgac ctgcgtcacg   31440
ttccgggcac gtcggtgccg gttgaatggc aggacgtgtc gggattgccg gaggacgctc   31500
agcacgaggc cgtgcagaag gtgctcgagg cggcgacgg cgtggcgttt gacgtggaac   31560
gcgggccgtt gctgcgcgtg accaccatga cgctcacccc cgagcatcac gtcttcgtca   31620
tggcggccca ccatctcgtg tgtgacggct ggtcgttcgg cgtcatcctg cgcgatctgg   31680
cgacgttcta ttcggcccag gttcgaggca cgcgtgcgaa cctcgaggcg ccgatgcagg   31740
tgagccgctt cgcgcgtgag gatcgcgagg ccaagcagag tcctgaagcg gccgagaccg   31800
aggcgttctg gatccggcaa ttcgacaccg tccccgagcc gctcgagtta cccgccgatc   31860
ggccgcggcc gccagtgaag tcgtaccagg gcgcccgtgt gtccgtgccg tttgacgcgg   31920
cgctcgctcg cggcgccaag aagctcgccg ccgaacaccg cacgaccctg tttacgacgc   31980
tgctcggcgt cttcgagacg ctcgtctacc gactgaccgg ccagcaggat ttcgtcgtcg   32040
gcatcccggt cgcggtgcag cctctgctcg gcgaggatct ggtcgcgcac tgcgtcaact   32100
tcgtgccgct tcgcgcgcgc gtgtcaccca ccgcgacgtt ctcggagtac ctgaacacgc   32160
tgaggacgca gtcactcgac ctcaatgagc accagaactt cacgtacggc agtttgctcg   32220
ccaagctgcg gttgccgaag caccgggcc gcgatccgct cgtctcgacg agcttcacgc   32280
tcgagcccgc catcgtcgat cccgggttcg acggtctcac ggcgcgttcg ctgacggttc   32340
cgcgcgtgac atcgaagcgc gacctgcacg tgaacgtcat ggagatggac ggcggtctgc   32400
tggtcgaggc cgactacagc accgatctgt tcgacgaacc gacggtgcgc cggtggctgg   32460
atagctatcg gatcctgctc gaggccgtcg tgtcgtcccc cggccggtcc ctcctgagcg   32520
tgccggtgat ctcggagccg gaccggagtc agcttgtcac cggctggaac gacacggccg   32580
ccgactaccc gcgcgatcgc gtcatccatc aactcgttga agagcgcgcc gtgcagaccc   32640
cggagggact cgcgctcgtc tgtgtggccg agcggctgac ctggaccgag ctcaaccgac   32700
gcgcaaaccg gttggcgcat cgcgctcgtga agaagggcgt cgcccccggg agcagggtgg   32760
cgctgtgcct ggaacgatcg gccgggttca tcgtctcggt gcttgccgtg cacaaggccg   32820
gcgccgcata tgtaccactc gacccgtct cgcccggcga tcgcaagtcc ttcgtcgtcg   32880
aggattccgg tgccgtgctg gtcgtgaccg attcgcggtc cgaaggccga tggctgacgc   32940
ccaggacgcc cgtgctccac ctcgacgccg acagtccgcg aatcgcgcaa gagtcgcacg   33000
atgatctgaa ggtcgccctg tcggccgaag acgccgcgta cgtgatctac acgtccggat   33060
cgacgggaca accgaagggc gtcgtcatcg ggcaccgca gctcgtgaat tacacgtggg   33120
cggtgatcga acggatcggg ttacccgccg gctcgagtta tgcgctcgtg tcgaacgtcg   33130
cggccgatag cggtgtgacc gggctctcgg cgctcgcgat gggctgggtg ctgcacgtga   33240
tcacctcaga ggtcgcgacc gacgggcggg cgctcgggtc gtatttgacc gttcacgaca   33300
tcgacttcct gaagatcacg ccgtcgctgc tcgcgtcgct gctcggcgat cacccgtcgg   33360
cctcgctgct gccgcgccgc tgtctcatgt tgggcggcga gccgtcacgt ccggcgtggg   33420
tggaagagct gcagcgtctc gcaccaggct gccggatcat gaaccactac ggcccgacgg   33480
agaccacggt cggcgtgttg acgtactgga ccggtgaccc cagggatctg ttcgtcgaga   33540
gcgtgccggc cggcaggccg ctcgggaacg tccgcgccta cgtgctcgat ccggccggag   33600
cgcccgcgcc gattggcgtc gtcggcgagc tctgtatcgg cggcgcgtgc gtggcacgcg   33660
ggtatctcaa tcgtcccgac ctcaccgccg agcgcttcgt ccccgatccg ttcgccaccg   33720
agccgggagc gcgcatgtac cggaccggcg accgcgccg gttccgggcc gacgggaaca   33780
tcgagttcct cggtcgcgcg gactatcagg taaaagtacg cggattccgc gtcgagctcg   33840
```

```
gcgaaatcga agcggcgctt cgcgctcatg agggcgtcga gcaggccgta gtcgtgctgc    33900
gaaaagatca gcccggcgac gagcgcctga tcggctatgt cacgaccggc ggcggcgcct    33960
cgatctcgat ggcggagctt cgcacgtcgc tcaagcagag gttgccggac tacatggtgac    34020
cggcggtgat cgtcgcgccc gacaagctgc cggtcacgtc gattggcaag atcgatcgac    34080
gcgccctgcc ggcccgcgcc gagcgcgcca gtttcgagtc ggagttcgtc acgcccgcca    34140
ccgagacgga gcgccggctt gcggagatct ggagcgcggt gctcgcggcc cgcaccgcca    34200
gcgcgcacga caacttcttc gacctgggcg ggcacctcgct gctcgcggcc cgcaccgcca    34260
tgaaggtccg cgacacgttt cgcgtggacg tgtcgctcat cacgttcttc gaaatgccca    34320
ccatcgccat gatggcggca gcaatcgacg ccggtaccgc ggccctgcc tcccggacca    34380
tcacgccgag ggcccgcgtc agggcccgca gggatgagct ggtccgccca tgaacccgg    34440
ccacccgcag actgagtccc ccgacctgcg gcgcgacggg atcgtcctgg cgcagctcga    34500
cgtgtttctg gcgccggcgt cgtacgccca ggggcgcttg tggttcctgg cgcagctcga    34560
gcctgaaagc tccgcctaca acattccgat cgcgctcaga ctgaagggcc gcctcgaccc    34620
cgccgcgctc gaacgcgcca tcaacaccat cgtccagcgt cacgagacgc tgcggacgac    34680
gttctcgctg cagggcgatg agctcaagca ggtcatctct ccgacgctgc gtattccgat    34740
ctccctcgtg gacatgcgtc tgctcagtcc tgagaaacgc gaggccggat ggcgccgcgt    34800
tgcggccgac gacggccgcgc gcgtcttcga tctggctgag ggcccgctgc tcaaggtcgt    34860
actcgtcatc ctcgagccct ccgatttcgt cctcctgatc acgtgcatc acatcgtctt    34920
cgacgggtgg tccggcggga tcttcatccg cgagctgccag tacgcggact ttgccgcgtt    34980
cgacgggcgc gacccgaagc tgccggagct gccggtgctg ttgaagcact ggacgcagga    35040
ccagaaggaa cagaccgaag ggcccggcct gccgacggat cgtccgcgtc cgcccctgca    35100
gctggcgggg gcgtcgatga cgctcgacct cctcgcttcg gctcgacgga gcgatctggc cgggcgtcgc    35160
gaccctcccgg ggccccctcg cctcgcttcg gctcgacgga gcgatctggc cgggcgtcgc    35220
cgcattcagc cggggcgaga acgcgacggt gttcatgacg cttctctccg ccttcggcgc    35280
cgtgctgcat cgctggtccg gacaggaaga catcctcatc ggcagcccgt cggccggccg    35340
cgatcgtccg gagcttgagg gcctgatcgg gttcttcatc aacacgttcg tcctgcggct    35400
caacgcgtcg ggagatccca cgtttcgcgga gttcgtcggc cgggcgcgcc gcgcgtgccg    35460
tgcggcctac gcgcatcagg acctgccgtt cgatcggctc gtcgagacgc tcaacccgga    35520
gcggacccgg gatcggcatc cgatctatca agtcatgttc gcgcaggatc cgccggcgtc    35580
gaagatgacc ttcgccggga tcgagctcga gcggcttgtc gtccacaacc gcagcaccaa    35640
gtgcgacctc acgctcgagg ttgccgaaga cgccgacgcg gtcaccctgt acgtcgagta    35700
cagcgtcgac ctgttcgatg cggcgacggt ggcgtcgctg ctgcagcaat atcgccgtgt    35760
gcttcagacc gcgatcgcgg acccggacca tcgcactggg gaccggtccg gcggccgtct accccggcga    35820
cgacgagcgc atgcagctga tcgcactggg gaccggtccg gcggccgtct accccggcga    35880
cgtcagtctg gccacgtgct tcgaagagcg ggtccgcgcg actccggcgg cagtggccgc    35940
gacgctggag ggacagcacc tcacgtacgc ggatctcaac cggcaggcga accgcctggc    36000
ccggcggttg aaggcgttcg gtgtgggccc gaacgtgctg gtcggcgtgg cgctcgaccg    36060
gtcgttcgat ctgctggtgg cgctcctcgc cgtggcgaag gccggcggcg cgtatctccc    36120
gctcgatttc gctcacccgc aggaacggct tgcgttcatg ctcgccgaca cgcgcgcgcc    36180
catcgtgctc acgctgcggc gatttgccgg ggcgctcgag tcctttcccc tcttctgcct    36240
ggacgacgag ctgccggcca ccgccgagga gccggacgaa gatctgccgc cgcagtcgac    36300
cggagaccac ctggcctatg tgatgtacac ctcgggctcg actggtcagc ccaagggcgt    36360
cgcgatcacg catcgcggcg tggttcgtct cgtacgcgttc gacgcgttga cgttcgaggt    36420
cgacgtgcgg gcgatcctcc acatcgcgcg gacgtcgttc cccgagcggc tcgtcagcct    36480
gtggggcgcg ttgctgaacg gcgcgcggct ggtgctggtt cccgagcggc tcgtcagcct    36540
ggagacgctc gagcgcacgc tccggtcgga gcaggtggac tgcctgctgc tcacgaccgc    36600
gctcttcaac gccgtcatcg acgcgaagcc cggcattctc gtcggcgtga agcaactcct    36660
gatcggcggc gaagcgctgt cagtcgcgca cgtgcggcgc gccctggcgc acctgccgga    36720
cacgcggctc gtgaacgcgt acggtccgac cgagtgcgtg accatcgcgt gcgcctacca    36780
gattccgcgc acgctcgatc cgcaggcgcg gtccattcg atcggccggc cgattgccaa    36840
tacgcaggcg ctcattctcg atcgacacgg cgacctcgtg ccgttcggca tcgccggcga    36900
gctctgcctc ggcggagacg ggctggcgcg ggaatacctg aatcagccgg cgctcactgc    36960
ggaacggttc gtgcccaacg cgttcagccc cgagccggcc gcccggctgt accgcacggg    37020
agacctcgcg cggatcgcc gcgacggcaa catcgagttc cttggccgga tggatcgcca    37080
gatcaaactg cgcggcttcc ggatcgagcc gggcgagatc gagcatgccc ttcgtgcccc    37140
gggagacgtt ctggacgcgg tggtggagat ccgggtcgat gcccacatgg cccatctcgt    37200
cgcctacgtc gtgcgcgccg atgggagtca actcaccgga accgacctgc gcgaacgtct    37260
gaaggcgcgc ctgccggaat actgtgtgcc cgcgaagtac gtgttcctgg ccgaggtccc    37320
gcgaacggcg gccggcaagg tcgatcgctc ggcgctgcct gcatcggcga tcgaggcgcc    37380
cgcgcccgag ccgcgcgatc cgtcgttcct gaacgaggtc gaagcggagc tcgcgcgaat    37440
ctggagcgaa gtgcttcggc tgccctccgt gctcgcgacc gacaattttt tcgatcgtgg    37500
```

//

Figure 5b

CDS   complement(3..914)
/blast_score=2e-66
/blastp_match="AE004644.PA2177 Pseudomonas aeruginosa"
/gene="regulator hybrid"
/note="probable similarity to prokaryote sensory
transduction proteins"
/product="sensor/response regulator hybrid"

CDS   924..2168
/note="none"
/gene="ligase"
/blastp_match="AP003013.MLR8297 Mesorhizobium loti"
/blast_score=e-152
/product="2-amino-3-ketobutyrate CoA ligase"

CDS   2207..3190
/blast_score=e-151
/blastp_match="AE008872.TDH Salmonella typhimurium"
/note="none"
/gene="dehydrogenase"
/product="threonine 3-dehydrogenase"
/pfam_match="PF00107; adh_zinc; 1"

CDS   3373..4455
/note="putative"
/gene="methyltransferase"
/blastp_match="AE001866.DR0026 Deinococcus radiodurans"
/blast_score=1e-08

CDS   4546..4959
/blast_score=2e-11
/blastp_match="AP002997.MLL1617 Mesorhizobium loti"
/gene="unknown"
/note="pfam00263, GSPII_III, Bacterial type II and III
secretion system protein, Expect = 7.8"

CDS   5176..6192
/blast_score=5e-98
/blastp_match="AF064070.PE20 Burkholderia pseudomallei"
/gene="glucose epimerase"
/note="putative "
/product="UDP-glucose 4-epimerase"
/pfam_match="PF01370; Epimerase; 1"

CDS   6331..14043
/note="substrat AT, malonyl ; zinc depend dehydrogenase ;
zinc dependent adenosine deaminase putative"
/gene="PKS I"
/blastp_match="AF285636.WCBR 2547 Burkholderia mallei"
/blast_score=0.0
/pfam_match="ketoacyl-synt"
/pfam_match="ketoacyl-synt_C"
/pfam_match="Acyl_transf"
/pfam_match="SAM binding"
/pfam_match="adh_zinc"
/pfam_match="pp-binding"

CDS   14275..15408
/blast_score=e-104
/blastp_match="AF285636.WCBT Burkholderia mallei"
/gene="acyl-CoA transferase WcbT"

```
                    /note="putative"
                    /pfam_match="PF00155; aminotran_1_2; 1"
CDS                 15436..16245
                    /blast_score=5e-11
                    /blastp_match="TTDEFFMT.FMT T.thermophilus"
                    /gene="formyltransferase"
                    /note="evidence experimental"
                    /product="methionyl-tRNA formyltransferase"
                    /pfam_match="PF00551; formyl_transf; 1"
                    /pfam_match="PF02911; formyl_transf_C; 1"
CDS                 16287..17384
                    /note="putative"
                    /gene="glycotransferase"
                    /blastp_match="AF285636.WCBD Burkholderia mallei"
                    /blast_score=e-99
CDS                 17427..18158
                    /blast_score=2e-82
                    /blastp_match="AF285636.WZM Burkholderia mallei"
                    /gene="ABC-2 transporter Wzm"
                    /note="putative"
                    /pfam_match="PF01061; ABC2_membrane; 1"
CDS                 18248..18847
                    /blast_score=7e-61
                    /blastp_match="AF285636.WZT Burkholderia mallei"
                    /gene="ABC-2 transporter Wzt"
                    /note="putative "
                    /pfam_match="PF00005; ABC_tran; 1"
CDS                 18952..20346
                    /note="putative"
                    /gene="glycosyltranferase"
                    /blast_score=e-101
                    /blastp_match="AF285636.WCBE Burkholderia mallei"
                    /pfam_match="Glycos_transf_1"
CDS                 20442..21167
                    /note="putative"
                    /gene="unknow"
                    /blast_score=1e-26
                    /blastp_match="AE009248.ATU3189 Agrobacterium tumefaciens"
CDS                 complement(21164..24301)
                    /note="tranporter domain"
                    /gene="unknow"
                    /blast_score=5e-29
                    /blastp_match="AE009122.ATU1658 Agrobacterium tumefaciens"
                    /prosite_match="PS00402; BPD_TRANSP_INN_MEMBR; UNKNOWN_1"
CDS                 complement(24351..27023)
                    /note="none"
                    /gene="unknow"
                    /blast_score=5e-29
                    /blastp_match="AP003581.ALR0267 Nostoc sp"
CDS                 complement(27806..29686)
                    /note="none"
                    /gene="cell surface protein"
                    /blast_score=2e-34
                    /blastp_match="AE010748.MA0851  2567 Methanosarcina
                    acetivorans"
                    /product="cell surface protein"
CDS                 complement(29535..30872)
```

```
                        /note="none"
                        /gene="cell surface protein"
                        /blast_score=2e-36
                        /blastp_match="AE010748.MA0851 Methanosarcina acetivorans"
        CDS             complement(30848..32647)
                        /note="fragment"
                        /gene="O-antigen"
                        /blast_score=1e-72
                        /blastp_match="AF105060.RFBC Riftia pachyptila"
                        /product="O-antigen biosynthesis protein"
                        /pfam_match="PF00535; Glycos_transf_2"
        CDS             complement(32574..35555)
                        /note="putative"
                        /gene="glycosyltransferase"
                        /blast_score=7e-46
                        /blastp_match="AE013462.MM2213 Methanosarcina mazei "
        CDS             complement(35533..36598)
                        /blast_score=7e-37
                        /blastp_match="AE013462.MM2213 Methanosarcina mazei"
                        /gene="glycosyltransferase"
                        /note="putative"
        CDS             complement(36516..37400)
                        /blast_score=8e-22
                        /blastp_match="AP003581.ALR0267 Nostoc sp"
                        /gene="ALR0267"
                        /note="putative ATP/GTP-binding protein, esterase fush
                        9e-14"
Sequence 37507 BP; 5531 A; 13507 C; 13011 G; 5458 T; 0 other;
```

```
    gatcccacga tcggcgccag cggattgcgc agctcgtgcg cgagcatcgc gaggaactcg        60
    gtcacgtggc ggccctcggc ctccagctcc tccatgcgct tgcggtgcga gaggtcgcgc       120
    gtcaccttgg cgaagcccat gtgccttcca ctcgcgtcgc gaagcgccgt cacgacgacg       180
    tcggcccaga agcgcgtgcc gtccttgcgc acgcgccagc ccgtgtcctc ggcgcggcgg       240
    tgcagcagcg catcgtgcag ctcggcctcc gggcggccgg ccgcggcatc ctcgggcgga       300
    tagaagagcg aaaagtgccg tccgatgatc tcttcggcgc tgtagccggt gatgcgctcg       360
    gcgcccttgt tccagctggt cacgtggccc gcggggtcca gcatgtagat cgcgtagtcc       420
    tgcacgccct cgatcaggag gcgcgagcgc tcctcgcttt gccgcagctg ctcctcgtgc       480
    tcgcggcgct cggtcaggtc gcgcgtgacc ttgccgaagc cgaccagcgt gccgtcctcg       540
    tcgcgcagcg ccgtgatgat gacgttggcc cagaagcgcc ggccgtcctt gcgcacgcgc       600
    cagccctcgt cctcgaagcg gccgtcctcg cgtgcgcgca gcagctcctt ggcggccag        650
    tcgcgcgcga tgtcctcctc ggtgtagaac gcggagaaat ggcggccgat gatctcttcg       720
    gcgcgatagc ccttcaactg ttcggcgccc gcgttccacg aacgcacgta gccctcggga       780
    tcgagcacga agatcgcgta gtccgtgacg gacccgacca gcagccggaa gatggcgtcc       840
    tcgccggcca gccgctcgg ggcgccgaag gcgaggttgc gcaccgaggg ccgcggaccg        900
    gatgggcttg gcatggctat tagggtacaa cggtgcacgc aagttgcaat cggacaatca       960
    gagataattg cttgatgaag atgccatttt tggaattctt acgcaggag ctggggacgc       1020
    tcgaggccca gggcctgtac aagtccgagc gcgtgctcga ctctccgcag ggcgccagcg      1080
    tgcaggtggc cgcccacgac gtgctcaact tctgcgccaa taactacctc ggcctggcca      1140
    accaccggc cctggtcgcg gccgcgaagg agacgctcga gcgcgacggc tacggcatgg      1200
    cgtcggttcg cttcatctgc gggacgcacg ccgtgcaccg cgagctggag acgcgcctcg      1260
    cgcgcttcct cggcaccgag gacgcgatcc tgtacggctc ctgcttcgac gccaacgtcg      1320
    ggctcttcga gacgctgctg ggcgagaggg acgcggtgat ctcggacgcg ctgaaccacg      1380
    cctcgatcat cgacgcgcgtg cgcctctcga aggcgaagcg cttccgctac gacaacaacg      1440
    acatggcgag cctcgaggcc cagctgaagg cggcggacgc ggcggcgcg cgctacaaga      1500
    tgatcgccac cgacggcgtg ttctcgatgg acggcatcgt cgcgaacctg ggcgccatct      1560
    gcgagctggc cggcgcgctac cacgcgatgg tgatggtgga cgactcgcac gccgtgggct      1620
```

```
tcatgggcgc gcacgggcgg ggaacgcccg agcattgcgg tgtcgaaggg aaggtggaca    1680
tcctcaccgg cacgctcggc aaggccctgg gcggcgcttc cggcggctac acggcgggca    1740
agcgcgaggt ggtggcctgg cttcgcaacc gctcgcgccc ctatctcttc tccaacacgc    1800
tgatgcccgc catcgcgggc gcgtcgctca aggtgctcga tctcctcgag ggcggcggcg    1860
agctgcgcgc gaagctcgcg cgcaacgccc gccacttccg cggcgagatg acgcgcctcg    1920
gcttcacgct ggccggcgcc gaccatccga tcatcccggt gatgctgggc gaggccccgc    1980
tcgcgaagga gatggcggac cggctgctga aggaaggcat ctacgtggtc ggcttctcgt    2040
ttcccgtggt gccccgaggg caggcgcgca tccgcacgca gatgtcggcc gcccatgaac    2100
cgaagcacgt cgatcgcgcc atcgccgcct ttgccaaagt cgggcgcgat ttgggagtca    2160
ttgcttgaga acccttcga agacgaagcg cgagcccggc atctggatgg tcgagtcgcc    2220
caagcccgtg gtcggccaca acgacgtgct gatccgcgtg aagaagaccg ccatctgcgg    2280
caccgacatg cacatcttca actgggacga ctggtcgcag aagaccatcc cggtgccgat    2340
gacggtcggc cacgagtacg taggcatggt ggaggccatg ggccaggagg tgcgcggcct    2400
gcaggtcggc cagcgcgtct ccggcgaggg ccacatcgtc tgcggccatt gccgcaactg    2460
ccgcgccggg cgccgccacc tgtgccgcaa cacgcagggc gtgggcgtga accgccccgg    2520
cgcgttcgcg gactacctgg tgattcccgc ggagaacgcg ttcccgattc ccgacgacat    2580
ccccgacgag atcgcctcga tcctcgatcc gttcggcaac gcggcgcaca ccgcgctctc    2640
gttcgacctg gtgggcgagg acgtgctgat caccggcgcc gggccgatcg gcatcatggc    2700
cgcggccatc gcgcgccacg tcggcgcgcg ccacgtcgtg atcacggaca tgaacgacta    2760
ccgcctggcg ctcgccacga agatgggtgc cagccgcgcg gtgaacgtat cgaaggaaaa    2820
cctgaaggac gtgatgcgcg agctcggcat ggtcgagggc ttcgacgtcg gcatggagat    2880
gtcgggcgtg ccctcggcct tccgccagat gctcgacacc atgaaccacg gcggcaagat    2940
cgccatgctc ggcatcccgc ccagcgaggc cgcgatcgac tggacccagg tgatcttcaa    3000
gggcctggtc atcaagggcg tctacgggcg cgagatgttc gagacctggt acaagatgat    3060
cgccatgctg cagagtggcc tcgacctctc gcccatggtc acgcaccgct tcgacgtgcg    3120
cgagtacctg aagggcttcg agacgatggg ctccgggaaa tccgggaaag tcgtcctgtc    3180
gtgggattag cgagggccac ggatcgcggc ccggttcccg cggcggcgat cgtgcaggcg    3240
ggttggcagc acgccccggg atttccgccg tggctcacgt gcgacttcgc cgtcgacgca    3300
ccgatcgcga tcgacgtcga cgtgagcgtg tcgctgctcg atgcccaggg cacctcgctg    3360
tggacgcagt cgatggccgg cagcgggcgc gacaccccgt ggctaccgac cggcacgtat    3420
cgcgcctccc tgcgcctgga gccgttccgc ttcccggccg cggccgcctc ggtcgagttc    3480
gcgctctcga tgcgggtgca aggtgagcgc cgcgtcgtgg ccaccgcagc gggcgtggtg    3540
cccgcgggcg cacccccacgg cgtggcgcgc gcggcctggc acctcgaggc cctggatggc    3600
acgccggcgc tcgagacgct ggcctggtcc gatcccacgc acagctggtt ctcccagcat    3660
ttcgaccatg cgtcgcgcac gtgcgtcgag tacctggggcg ccggatggcgc cgggtggcgc    3720
gggcgcgtgc tggatgcggg atgcggcgac ggcatcacgg cgctcggcat cgcgctgcgc    3780
tatgcgcccg agcaggtggt gggcatcgat ccggggcgct gctaccgggt gcttcccgac    3840
atcctcgagc gccacgccct cggccacctc gcgctgccgg acaacctgca gttcctcccc    3900
gcggatgcca acgcgctccc gttcgccgac agcagcttcg acgtcatcgt ctcgtggtcc    3960
gcggtcgagc atttcgtggg cggctacctg ccctcgctcg ccgaggcgcg gcgggtgctc    4020
aagcccggcg gcctgctggt catccatccc gagctctact acaccgcgca ccacggccat    4080
cacctcggcg agtacagccg cgagcccttc ttccacctgg tgaagtcgcc cgacgaggtg    4140
cgcgacatcg tgttcggcgc ggacgtgagc ctccacgacc gcggcggcat cgcgccgacg    4200
cgcgcggagc actggcgctg gtacaacgag ctcaatcgca tcacggccgc gcagttcgag    4260
gacgagttgc gcgcgctcga cttcgagccg tggcgcctgg cactgcgcgc cgagccgctg    4320
gtggagtaca agccggagct gctgcgctac cgcttcacgg agctgggcgt ctccgagctc    4380
tacgtggcct gcatcaaccg caagcccgtg tcagatggtt cctacaggac gcatgggcag    4440
gcggcattcc ggtaaggtcc gccgggacgg gaattgaagg cgcaagcgcg ttgttatccc    4500
cgtcactccc ggactgccgg gatcaccctc acaggagaac ccccgatgaa gacccagacc    4560
caggccgtgc tcgccgcggc cctgctgtcc cttgccggca ccacgctggc ccgcgccgaa    4620
gccgtcacgc tgacggcgc gagcgaagtc ccgcccgtca cctcgagcgc gaagggcagc    4680
ggcacggtgg tcgtgaaggc cgattgcacg gtcaccgcga agatcaccgt caccggcatg    4740
accgccaccg cggcccacat ccacgaaggc aaggccggcg ccaacgcggg cgtggcagtc    4800
ccgttcgtga agaccgcgga cacacggttc gaggcagccc cgggcgccaa gatgaccgag    4860
gcgcagtgcg ccgcgtacaa ggcgggcggc acctacgtga acgtgcacag cgagcccac    4920
aagggcggcg aggtccgcgc ccagctggcc ggcaagtagg gcagccgctg caaaaagaaa    4980
aagcccgcgc aagcgggctt tttcgttgaa ggctttggtc ggggcgaaag gattcgaacc    5040
```

```
ttcgaccccc tgcacccccat gcaggtgcgc taccaggctg cgctacgccc cgaccgaact    5100
tgaaattata cctgcatcgg ggtcgcagcc gccatgaatc gaaacgcccc gggcttcggg    5160
tagcatccgc ccatggcgac gaccctcgtg acgggtggtg caggctacat cggcacccac    5220
atcctctgcg cactggccca ggccgggcga cgcagcatct gcatcgacaa ctactccaac    5280
agctcgccgc gttccatcga acgcgtgctg cagatcgcac ccggttgcgt cgaggccttc    5340
gacgtggaca tccgcgacgc cgacggcatc cgcaaggtga tcgcagggcg cgacgtggac    5400
agcgtgatcc acctcgcggg actgaaggcg gtcggcgaat cggtcgagca gcccgagcgc    5460
taccacgaca acaacgtccg cggcaccgag agcctgctcg cggcgctcgc cgactcgccg    5520
gtgcgcaagt tcgtcttcag ctcctcggcc accgtgtacg gcctcgccga gaagatgccg    5580
atcgacgagg acgcgcccac ctcgccgcag agcccgtacg gccagaacaa gctcgacatc    5640
gagcacatgc tggtcgccct cgccaagcgc gacccctcct ggcgcgtagt gaacctgcgc    5700
tacttcaacc cggtcggcgc acacgagagc gcgctgatcg gcgaggaccc ggccggggtt    5760
cccaacaacc tcatgcccta cgtgtgccag gtcgcctccg ggcgactgaa ggagctgagc    5820
gtctacggca gcgactaccc cacccccggac ggcaccggca tgcgcgattt catccacgtc    5880
tgcgacctgg ccgaaggcca cgtggccgcg ctggaagcac tcgagcgcgc ggcccccggg    5940
accgtcctca cggtgaatct cgggacgggg ctcggctaca ccgtgctgca attgctcgaa    6000
accttcgagc gcgtgaacaa actaagcatt gcccggcgca tcgtggggcg gcggccgggc    6060
gacgtcgcga tctgctacgc gaacgcgggc cgcgcgaagg ccgtgctggg ttggaccgcg    6120
aagcgcggaa tcgaggagat gtgccgcgat gcctggcgtt ggcaggaaag gaatgcaaag    6180
ggctacggtt gacggattgc aaatcatggc cttggcgacc cttgctgcaa tgcggcatga    6240
agggggcggt tcttgtaaga tccaaggttc gggatgcgtc caggacgtca aaaacagcac    6300
ccggcgttgg aaggaagaag gcaaccctag tgttgtcatg gctccagacg ggattgcgcg    6360
ccgagtaaga atcgtcagtc gagaccaact cgcctccggg tgcaacgaat cgtgagcccga    6420
agcaacacct cgcatcgcgt cgccgtcgtc ggcattgcgt tccgccttcc cggcgctacg    6480
cgcgaatccc tctggccctc gctcctcgag gggcgcaacc tcgtcaccac cgtcgatccc    6540
gcgcgctggg cgaaggagcc gtactaccac ccgcgcaagt ccgagcccgg caccagctac    6600
tcgttcgccg ccggcaccat cggcgacatc agcggcttcg acgcgtcgtt cttcggcatc    6660
tccccgcgcg aggccgcgca gatggatccg cagcagcgcc tgctgctcga gctctcctgg    6720
gaagccctcg aggccgcggc ggtgaagccc tcgaccctgc gcggcagcga ttgcggcgtg    6780
ttcgtgggcc tgtccaccgt cgactacagc tacggcctcg cctacgacct ggcgtccatg    6840
gacgcgtcgt cggccaccgg caacaccgcg agcatcgcgg ccaaccgcct ctcgtatttc    6900
tacgacctcc acggcccccag catggtcgtg gatacggcgt gctcgtcctc gctggtggcg    6960
ttccaccagg cctgccagtc gatcgccatc ggcgagacgc gccaggcgct ggtgggcggc    7020
atcagcctgc acctgcatcc gtacgggttc atcgcgttct ccaaggcgtc gatgctctcg    7080
aagcgcggcg cctgccgcgt gttcgacgcc agcggcgatg gctacgtgcg ctcggaaggc    7140
gcgggcgtgc tggtgctgaa ggacctcgac cgcgccgtgg ccgacggcaa cccgatcctc    7200
gccgtcgtcg ccgcgagcgg cgtgaatacc gcgggccgct cctccagcct caccgtgccc    7260
agcgtcgatg cccaggctgc attgctgtcg gatgtgtacg agcgcgccgg catcgacccg    7320
tccgagatcg actacctcga ggcccacgac accggcaccg tggtgggcga tccgatcgag    7380
acccgcgccc tcggcatcgc gctcgggcgc ctgcgcccgc gccaccagcc gctgccgatc    7440
ggttcggtga agagcaacgt cggccacctc gaggcggcct cgggcatggc gggcctggtg    7500
aaggcgctgc attgcctcaa gcaccgcgcg gtgccgccca ccatccacct cgattcgccc    7560
aaccccaaca tccacttcga cgagtggaac ctgaagccgg tcctcgagcc gctcacgctc    7620
catccgtaca agaagctcgt cgtcggcgtg aactcgttcg gcttcggcgg cgccaacgcg    7680
cacgtgatcc tcgagagcac gcccacttcg cccaaggcgg tcgaggcccc ggccggcgag    7740
gtggcgctca tgctgagcgg ccacgacgac gcggcgctgc gcgacgtggc ggcctcgtac    7800
gcgaacttcc tgcaggcgag cgagcaaccc gcgctgcacg acatcgcgta caccaccgtg    7860
cactgccgcg actggcacca gctccgcgcc atggtggtcg gcaccgaccg cgccagcatc    7920
gccggcgccc tcgcgcaatt cgccaagggc cacaacacgc ccggcgtcgt gagcggccgt    7980
gcgctgccca acgcgtccgg ccccgcattc gtgtattcgg gcaacggcgc gcaatggccc    8040
gcaatggggc gcgccctgct cgccggcgag ccggtgttcc gtgccgccgt gcaatcggtc    8100
gacgcgctgt tccgcccgct cggcggctac tccatcctcg agatgctgga gcaggaggtc    8160
cccgccgacc agctggaccg caccgagatc gcccagccgc tgctcttcgc cgtgcaggtg    8220
ggcttgaccg agctgctgcg ccactggggc atccagccca cgcgcgtcac cggccacagc    8280
gtgggcgagg tcgccgccgc ttgggcgagc ggcgcgctga cgctcgacga tgcggcgcgc    8340
gtgatctacc accgcagccg ctaccagggc gagaccaagg cgcgggcgc gatgaccgcc    8400
gtcggcctcg ccgaggacgc cgcgcgccgc cagatcgagg cgctcgggtt gaagggcgag    8460
```

```
gtgaacgtcg cgtgcatcaa cagcacgcgc aacgtgacgc tcgccggcac gcgcgccggc      8520
atcgaggcgc tcgaggccga gctcacgcag cgcaaggtgt tccaccgccg cctcgacctc      8580
gactacgcgt tccacagccc ggcgatggac ccggtgcgcg acgggctcgt gagcgcgctg      8640
cgcggcctca cgccgcgcgc cacccgggtc gcgttccatt cggcggtcac cggctcgccc      8700
gccgcgggca accagctcga cgccacctac tggtggagca acatccgcga gccggtgcgc      8760
ttccaggccg cgatccgcgg catcgccgag tcgggcgtca atgtcttcat cgagatcggc      8820
ccgcacccga tcctcaagaa ctacatcaac gacggcctgc gcgccgcctc gatcgagggc      8880
cgcgcgctgg tcacgctgca gcgctcggcg agcgatcgcg cggcgatccg cgccgccgcc      8940
caggaagtcg tgatcaccgg ctgcccggtg gagaccgcga agatgctccc cgcgcagggc      9000
cacttcgtcg agctgccgcc ctacccgtgg cagcgcgagc gccactggcg cgcgcccacc      9060
tcgcaggcct acgacctcat ccagcacggc aagcagcacc cgctgctggg ctatcgcctg      9120
cacgagaacg acttccagtg ggaaaaccac atcgacaccg cgctctatcc ggcgtacgcg      9180
gaccatgtcg tcggcggcgc cgtggtgttc cccgcggccg gcttcgtcga gatggcgctc      9240
gccgcctcgg cgatcgcgct gggcggcgac gcgcacgaga tcgagacgct cgagatccgc      9300
agcgcgctgc tgctcgagga ctccacgtcg aagaccgtgc gcttcgcgct ggaacccgat      9360
ggccgcttca caatccgcag ccgcgctcgc ctgagcgagg atccgtggca actccacgtg      9420
gtcggcaagc tggtcggcac gcccaccgag cttccgcgcg cgccgcgcca gtcggcgccg      9480
tcgcgcaagc ccgatgtgat cgcggcccag cactacgaag gcgccgcgaa ggccggcctc      9540
gcctacggcc cggccttcca gtcggtctcc aaggtgtggc tcgacgccgc ggacccgggc      9600
agcgcctacg cgcgcctgat cctccccaag ccgattcgcg cgcagctggg cgtgatgcac      9660
ctgcaccccg cgtcgctgga cggctgcttc cagctgctgg tcgacctgct gcgcgccgag      9720
gcctcccggc acgcgcagat cgccttcgtg ccgatcctgg tcgggcgcac gcgcctctac      9780
ggccctgccg ccctcgtcac cgcctgccgc gtccgcctga ccgcgcgcag cccgcgctcg      9840
ctggtcgcgg acttcgagct ctacggcccc aacgcgcacg tggtggccgc gctcaccggc      9900
gtgcgcttcc gcggcgtcca gctgaagcag ccgcagggcg cgcgcctgcg ctacctgtcg      9960
taccgcgcca tcccgcgcct gaatggcggc gacacgccgg cgcgccaagcc gctgcccctc     10020
gcgttgctgg ccgcggcgtg cggcgagcgc tggcattcgc tctcggccag ctcgcgcaag     10080
cgctactacg atgaggtcga gccgctctcc gacgtgctgt gcagctcctt cgccgagcac     10140
gcgctgcgaa agcggatcgg cgaggccccg ctcatcgacc cggccgcgct cgccgcgtcg     10200
acgccggccg agcattggcc gctcctcaag cgcgcgatcc agatgctggt cgaggaccag     10260
ttgctcgagc ccgccgaagg cggctggcgc tggagcccgg tcaccgaatt gcccgacgcg     10320
caggagagct ggatcacgct gctgcgcgac tacccggacg aggccggcca gctgatcgcg     10380
ctcgggcgcg ccggcgctca cctcgcggaa gtgttcgccg gcaaggaagc gtccttgccg     10440
agcctgttgc cggccacgca ggcgaacgtg ttcgcgcgcc tgtgctcggg cgcacccgcg     10500
ttcgccaacg ccgggcgcgc catcgccgac acgctggcgc tcgccctcga gcgcctgccc     10560
gcgcaccggc cgctgcgcgt gctcgaggtg acgccgtgcc gctccgagat cacgctgcgc     10620
gtcctcgagg ccgtggacct cgaccgctgc gagtaccacg tcggctgcac cacggacgaa     10680
gcactggcg agtacgaagg cgtcctcgat gccattccgc acgtcgagac ttgcgtcgtg     10740
gacctgaaga tcccgggcat gggcctgaag gcgccgaacc agggcccgtt cgatgtcgtg     10800
atcgtttccg atggcctgct cggtgcggcc gatccggatg cggccgcggc gcacctcgcc     10860
ggcgcgctcg cggaagacgg gctgctggtc atggtcgcgc agcatccgtc gcgctggacc     10920
gacctcctct tcggcgcgga cccggcgtgg tggacgctcg ggccgctcgg ctcgcagcgc     10980
tcgcgcctgc gttcgccgga cgagtggcgc atcgcgctgg cccacctcgg cttcgggccc     11040
accgtcgcgg tgcccgagct gccggccctc gagcgcggct cgtacgtcct catggcgcgg     11100
cgcgacggcc cggtgccgca ggctgccgag cgcagcgttc ccgccaagtc gggctggatc     11160
ttcgtgcagg aagcgggcgg ttattcggcg gcgttcggcg cctgcgtgca ggcggatctg     11220
cgcgcacgcg gccacgcggt gttcgagatg gcgcccgatt ccgtcgaggg ggccgagatg     11280
gcgcagcttg cgcgcgcgcg ccagacgctg ggcgacatca gcggcatcgc cttcctggcc     11340
ggcctgcccg acgacgcggc cgccacgggc gacgcggccg agcgcctcgc gcgccagacc     11400
gcgcgctgcg ccgcgctcgc gcgcctgctg cgcgaatgca aggcgagtgc catgacaccc     11460
gcggtgtgga tcgtgacctc gggcgcggcc tccagcctgg cgcgcaacgt ccccgccctt     11520
gcctcgcgcc gccgccctcc cgatccggac gaggcgatcg tatgggggtt cgcgcgcacc     11580
gcgatgaacg aattccccga gctgcgcctg cgcctcgtgg actgcccgaa cccggagaac     11640
ctcgagcgca acgccgcctc gctggtgcag gagctgctct cgcccaccgc cgacgacgag     11700
gtggtcctca ccggcgccgg ccgctacgcg ctgcgcctgg gcgcggtgcc gccgcccgca     11760
tccgtgcgca ccgctaccca gggcatcacg cgcctcgact tcgccgcgcc cggccgctg      11820
aagaacctcg cgtggcgcgg cgacgtgcgc cgcaagccgc gcgcgaacga ggtggaaatc     11880
```

27/39

```
gaggtgcgcg  ccgcgggcct  caacttccgc  gacgtgatgt  acgcgatggg  cctgctctcc      11940
gacgaggcgg  tcgagggcgg  cttcgccggc  gcctcgctcg  gcatggagct  ctcgggcgtc      12000
gtcaccgccg  tcggcaagga  cgtgacctcg  gtcgcgcgcg  gcgacgaggt  gctcgcgttc      12060
gcgccctcgg  cgttcgccac  gcacgtggtc  accaccgccg  attcggtcgc  gaagaagccc      12120
gccggctgga  cgttcgagtc  cgcggccacc  atcccgaccg  cgttcttcac  cgtgtactac      12180
tcgttgaagc  acctggccca  gctgcgcgag  ggcgagcgcg  tcctgatcca  cggcgccgcg      12240
ggcggcgtgg  gcatcgccgc  gatccaggtc  gccaagtggc  tgggcgcgga  gatcttcgcc      12300
accgccggct  ccgacgagaa  gcgcgacttc  gtgcgcctgc  tgggcgcgga  ccacgtgctc      12360
gactcgcgca  cgctcacgtt  cgccgacgac  gtgctgcgca  tcaccaacgg  cgagggcgtc      12420
gacgtggtgc  tgaactcgct  ctccggcgag  gccattgcgc  gcaacctgcg  cgcgctgcgc      12480
ccgttcggcc  gcttcatcga  gctcggcaag  cgcgactact  acgagaacac  gcacatcggc      12540
ctgcgcccgt  tccgcaacaa  catctcgtac  ttcggcgtcg  atgccgacca  gctcatgaag      12600
gagcggcccg  acctcgcgcg  ccagctcttc  aacgagctga  tgcagctctt  cgagcagggc      12660
gtgctctcgc  cgctgcccta  ccgcgccttc  gccgccaccg  aggcggtgga  agccttccgc      12720
tacatgcagc  actcgcgcca  gatcggcaag  gtggtgctgt  cgttcgccga  cggcgtgaag      12780
gcgcagcccg  cccggcccgc  cgtgccgcgc  gagctggccg  tgtcgccgaa  cgccacgtac      12840
ctcgtcaccg  gcggcctgtc  gggcttcggc  ctgcgcaccg  cgcagtggct  ggtggacaag      12900
ggtgcccgcc  acctggtgct  ggtctccaag  agcggcgccg  agtcgctcga  gaccaaggcc      12960
gcggtggccg  acctcgaggc  ccagggcgcc  acggtgatcg  cggaggcgtg  cgacatcacc      13020
gaccgcgcct  ccgtgcaacg  gctgctcgcc  gaggtcgcgg  ccgcgctgcc  gccgctgcgc      13080
ggcgtgatcc  acgcggcggc  cgtgatccag  gacggcttca  tcaccaacat  gacggccgcg      13140
cagatccgcg  acgtgctggc  gcccaaggtc  ctcggcgccc  gccacctgga  cgagctcacg      13200
cgcggcgcca  agctcgattt  cttcgtgctg  tactcgtcgg  cgacgacgct  gttcggcaac      13260
ccgggccagg  ccaactacat  cgcggccaac  tgcttcctcg  aggcactcgc  caaggcgcgc      13320
cgcgtcgagg  gccagccggc  gctgtgcgtg  ggctggggcg  cgatcgggga  cgtgggctac      13380
ctcgcgcgcc  acgagaaggt  gaaggaagcc  ctgcagacgc  acatgggcgg  caccgcactc      13440
gagtcggagg  ccgcgctcgc  ggtgctggag  cagctcctgc  tcgccgacgc  ctcgggcctc      13500
ggcgtgctcg  acttcagctg  gcgcaacctc  cgccgcttcc  tccccaccgc  cgcctcgccg      13560
cgcttccgcg  agctgggcgc  gcgcggcgac  gacgcgaagg  acgacgacga  ccggcgcggc      13620
gagctgcgcc  gcctccgcca  cgagctggac  gcctccgagc  tctcggcgat  gttcaccgac      13680
ctgctgcgcc  gcgaggtggg  cgagatcctg  cgcatcccgc  cggaccgcct  cgatacgcgc      13740
cagcccttgc  aggaaatggg  catggattcg  ctgatgggcg  tggagctgct  gaccgcggtg      13800
gaggcgcgct  tcggcgtgaa  cctgccggtg  atgtcgctct  ccgagcagcc  gtcgatcgag      13860
aagctggtcg  accgcatcgt  gcgcgcgctg  aaggacccga  acgccggcac  cgaggccgaa      13920
tcgcgcagcg  accagatcga  gcgcgtcgcc  gcgcagtacg  cgcccgagct  cgactcgcgc      13980
caggtcgagg  agctcacgca  ggcggtcgat  gacgcgcaag  ccacgcagca  gggacgtccg      14040
tgaaggcgcg  caacctccgc  ggcctctcgc  aggcagtgaa  ggagcagctg  atccagcgcg      14100
tgctcgagca  gcgcgtgcgc  cgggtcgaga  aggacaagcc  ctcggcggtc  gagtccgcca      14160
tcaacgcctt  ccgccccgac  cccgatgccc  tgggccacat  ctcggagcgc  tactgccgct      14220
tcgacatgca  cccggtctac  cagcagatgc  agctggtgaa  ggaaggtgcc  gcgaagctgg      14280
gcattcccga  tccgtacttc  cgcgcgcacg  acggcgtggc  cgcggatacc  acgcgcattg      14340
ccgggcgcga  gtacatcaac  ttctccagct  acaactacct  cggcttctgc  ggcgacccgg      14400
ccgtggatgc  cgcggccaag  caggccatcg  acgaatacgg  gaccctcggt  ctcggccagc      14460
aggctcgtgt  ccggcgagcg  gccgctgcac  cgcgagctgg  agcgcgccat  cgccgaggtc      14520
tatggcgtgc  acgacgccgt  ggccttcatc  agcggccacg  cgaccaacgt  ctccaccatc      14580
ggccacctgt  tcgggccgcg  cgacctcatc  gtgcacgacg  cgttcgtgca  caacagcatc      14640
ctgatgggca  tccagctctc  gggcgccaag  cgcatggcct  tcccgcacaa  cgactggcgg      14700
gccctggacg  agctgctgcg  cgcgcagcgc  cgccacttcg  agcgcgtgct  gatcgtgatc      14760
gagggcgtct  acagcatgga  cggggactac  cccgagctgc  ccgaattctc  gcgcctgcgc      14820
cgccgccacc  gcgcgttcct  gatggtggat  gaagcgcact  ccttcggcgt  gatgggcccg      14880
cgcggcttcg  gcatccgcga  ccacttcggc  atggccggcg  acgaggccga  catctggatg      14940
ggcacgctct  ccaagacgct  cgcctcgtgc  ggcggcttca  tcgccggcga  gaaggcgctg      15000
gtcgagcacc  tgaagttcgc  cgcccccggc  ttcctctaca  gcgtcggcat  ggccccgccc      15060
gtggcggccg  cggcgctcgc  cgcgctgaac  cgcatgcgcg  aggagccggc  gcgcgtggcc      15120
gccctgcagg  cgcgcggacg  cttcttcctc  gaggccgcgc  gcgccgcggg  catcgacgtg      15180
ggcctctcgc  agggcatggc  cgtggtgccc  gccatcaccg  gcagctcgat  ccgcgccggg      15240
cgcctggccg  aggcgatgtt  ccagcgcggc  atcaacgtcc  agccgatcgt  ctaccccggcc     15300
```

WO 2004/013327   PCT/EP2003/007765

```
gtgcaggaga actcggcgcg cctgcgcttc ttcgtgagcg ccacgcacac cgaggagcag    15360
ctgcgcttca cggtccgcga gctcgccgac gcctggcgca agctctgagt ggcgggcccc    15420
cggctgcgcg tcgtcgtctg cacccacggc ggcctgaacg cgcgctggt gctctcgcgg     15480
ctgctggccg cccccacgct cgaagtctcc gcgctggtga tctccagccg cgcgcgcggt    15540
gcccatgagt cgtatggccg tgccgcgctc ggatacgtgc gcgcgagcgg cgtcgcctat    15600
gcgctgtacc tctggtgcgc cacggcgctc gccgacctgc tgctgcgcgg cacgtccgag    15660
ggcccggtcg cgcgcatcgc cctcgcgcgc ggcatcccgc tcctcgccac gccgcgcgtg    15720
aacgatgcca ccgcgcgcgc cttcatcgcg ggggcgcgc cggacctgat cgtctccgcg     15780
ttcttcaacc agcacatcga cgccgacgtg gctgcgctcg cgcgcgtggc cgcggtcaac    15840
atccacccgt cgccgctgcc gcacttccgc ggcgtggatc cggtgtcctt cgcgcgcctg    15900
cgcggtgccg agcgccacgg cgtgagcgtg catcgcatcg aacccggctt cgataccggc    15960
gcgctgctcg cccaggaaac cgacgtcgag gcccccggca gcgtctttgc cgccaccgcg    16020
gcgctctatg accgcgggggc ggcgctcctc gccggccgtg ccgcggcgct ggccgccgac    16080
ccgcgcggaa cccccagcc ggcggggggc tcctacgact cctggcccac ccgcgcccag     16140
gtcgccgcct tccggcgcgc cggcggggcgc ctgctccgcg cccgcgacct gtggcgcctc    16200
gcgcgccggg ggccggccgc tttcgtaata gaatcagcgc ggtagcggcg actcacgccg    16260
cgcttttccc cgacccccccg aggcccatga aacactggct gaagcaacac gcaatcttcg    16320
cgctcaccgt cctgctgccg accgtggccg cgatcctcta cttcggcctg atcgcctccg    16380
acgtctacat ctccgaatca cgcttcgtgg tgaggagccc ccagcgccag gtgcagaccg    16440
gcctggtggg cgccctgctc tcgggcaccg gcttctcgcg ctcccaggac gacacctact    16500
cggtgcacga cttcatcacc tcgcgcgacg cgctgggcga gctggacaag aagctcgccg    16560
tccgcaagct gccaacatcg acttcatcaa ccgcttcccg gggctcgact    16620
gggacgacag cttcgaggcg ttccaccgct actaccagaa gaaggtcacg atcgacttcg    16680
acaccgcgtc ctcgatcacg gtgctgcgcg tgcgcgcctt cgagaaggcc gactcgcggc    16740
gcatcaacga cctgctgctg cagatgggcg agcgcctggt gaacgagctg aacgagcgca    16800
gccgccagga cctgatccgc ttcgcgcagg ccgaggtctc gctcgccgag gacaaggtga    16860
aggatgccgc gctggcgctc tccgccttcc gcagcaacca gtcggtgttc gagcccgacc    16920
gccaggcctc gatccagctg cagggcgtgg ccaagctcca ggaggagctc atcgccaccg     16980
agggcagct ggcacagctg cgcaagctct cgcccgacaa cccgcagatc ggcgcgctcg      17040
agaacaagtc ggcggcgctg cgcgtggcga tggcgcgcga gtccgccaag gtgaccggcg    17100
gcagcggctc gttcagcgcg cgcgccccgg cgttcgagcg cctcaccctc gagaagggct    17160
tcgccgaccg ccagctgggc gttgccctca ccgcgctgga gaccgcgcgc agcgaggcgc    17220
agaggaagca gctctacctc gagcgcatcg tgcagcccaa cctgcccgac gaatcggtcg    17280
agccgcgccg catccgctcg atcttcaccg tgttcgtgct gggcctggtg gcctggggcg    17340
tggtgagcct gctggtggcg agcgtgcgcg agcacgtcga ctaggccgtg gagcagaacc    17400
cttccctcct gcgcgcgcgc ggggtcgacg ggcgcgtgct ctacgcgcg ctgatgcgcg      17460
aggtcatcac gcgcttcggg cgcgacgacc tcggcgtgct gtggctggtg gtcgagccga    17520
tgatcttcac gctcggcgtg acggccctgt ggaccgcggc cggcatgaac cacggctcgt    17580
cgctgccgat cgtggccttc gcggtcaccg gatactccgc ggtgctggta tggcgcaact    17640
gcgcctcgcg cgcctcgatg gccatcgagg ccaacaaggg cctgctcttc caccgcaacg    17700
tgcgcgtgat cgacgtgttc gtgacgcgca tcatcctcga gatcgcgggc gccacggcgt    17760
cgttcaccgt gctgggcatc ttcttcatct ggatcggctg gatgccgatg ccgatcgaca    17820
tgctcaaggt cgcgttcggc tggttcatgc tgtcgtcggtt cggcgccagc ctggcgctcg    17880
ccatcggcgc cggcaccgcg tactcgctgg tggtcgagcg gctgtggccg ccgacggcct    17940
acctgctgtt cccgctctcg ggcgccgcct tcatggtcga gtggctgccg gagaaattcc    18000
agcagttcgt gctcttgctg ccgatggtgc acggcaccga gctgctgcgc gacggctact    18060
tcggcaacgc ggtgcgcacg cactacgacg tgggctacat ggcaacggtc tcgctgggcc    18120
tgtcgctgct cggcatgcac ctggtgcgca tggcggccaa gcgcgtggag gccgcgtgat    18180
ccgggtcgag aacgtgacca gatgtactca cgcgcctcgg cccagaacgt gctgacgcgt    18240
cagcttcgag ctcgacgtcg ggcgcaacat cggcatcctc ggcaagaacg cgcgccggcaa    18300
gtccacgctg atccgcctga tcagcggcgc cgagagcccc accaccgggc gcatcgtgcg     18360
cgagatgagc gtgtcctggc cgctcgcctt cggcggcgcg ttccagaccc acctcaccgg    18420
gctggacaac ctgaagttcg tctgccgcat ctacggcgtc gactaccgcg acaaggtcgc    18480
cttcgtcgag gacttcaccg agctcggccgt gtatttccgc gagccggtgt tccactactc    18540
gcacggcatg atcacgcgcc tcgccttcgc gctgtcgatg gcggtggagt tcgactgctt    18600
cctgatcgac gaggcgatgg tggtgggcga cgcgcgcttc cacgagcgct gccacgtcga    18660
gctcttccac aagcgcaagg accgcgcctt catcctcgtc acccacgacg ccaaggtgat    18720
```

29/39

```
caagctctat tgcgagagcg cctgcgtgct gcacgagggg cggctgctgc ccttccccac   18780
cgtggacgcc gcctacgagt tctacatgaa cgaggtcatg caggacctcg ccccggaggt   18840
tgcctgaacg cccccgcgcg gccgctgcgc gtgctcctcg agctgcggcc ggcgttcgac   18900
ggccacgccg gcatcccgca ggaggcacgc ctgctgttcc gcggcctgcg catgatcgag   18960
ggcatcgacg tggagggcct gctgcagcac agcggccacg tgctcgccaa gggcctgccg   19020
ccgcgcggcg gcggcgacct cgcaccggac cgccagctga accgcctgtc gcgcgtcgtg   19080
gtctccacca agcaggagct caccaacgcg cacgtggcga cggcgctgat ggccattcgc   19140
cagctgctgg gcgcgcgcga gaccctctac cgcttcgacg cggcgcactt ccgcgacttc   19200
gtctggcagg cgctcttcgc gcgcacgctg cacgccgcgg acttcgactc ggtcacccgc   19260
gccggctttc gcgtcgcgcg cgtgccgtcg acgggcatgc accgctgcgc gctggtcacg   19320
cgcaagctcg gctacacgct gcatccgcgc atcgacactt ccgatttcga cgtgctgatc   19380
gccgagacgc cctacccggc gcgggtcgcg gacggcacgc gcctcgtggt gcgctaccac   19440
gacgcgatcc cgctcctcat gccgcacacc atctccgaca tgtcctacca ccaggcctcg   19500
cactaccggg cgctgcgccg caacgtggcc tcgggcgcgc acttcgcgtg cgtctccgaa   19560
gccacgcgca aggacctgct ctcggtcttt cccgaggtcg aggcgcgctc cagcaccatc   19620
cacaacatgg tctcgcacca ctacttccgc gagaccgccg acgccgggcg catcgaggag   19680
atcctgcgaa cgcgcgcgag cgagcgcatc aagggcgcgc ccaacggagc cgccaccgcc   19740
gtcgcacgat cgccctcctt cctcgcgcgc gccgcgcaac ccagcgaccc cggctacctg   19800
ctcgccgtgg cgacgatcga gccgcgcaag aaccacgcgg gactgatcgc cgcgtgggag   19860
cacctgcgca cctcgcgctt cccgggcctg cgcctggtgg tggtcggcat gcccggctgg   19920
ggctacgagg cgctggtagc caagttcaag ccgtggctcg cgcgcggcca gctcttcgtg   19980
ctcgaggacg tgccccgcccc cgagctgcgc ctgctctact cgcacgcgcg tgccacggtg   20040
tgcccgagct tcggcgaggg cttcgatttc tccggcgtgg aggcgatgcg ctgcggcagc   20100
ccggtgatcg cctcggagat cgccgcgcac cgcgcgaggtgt accgcgacgc cgccgagtac   20160
tgcagcccct attccgtggc cgacctcgcc gaggcgatcg ggcgcgtcat cgacccggcg   20220
gcgacgggcc tgcgccaggc gctggtcacg cgcggcaccg aggtctcgca gcggtacacg   20280
cccgaagcca tcctgccgca gtggcgcgag tacctgctcg gcacggtgcg cgcggaggcc   20340
ccgtgagcga tccgaaggcc accgaccagg cgtggggagga gtggggccag cgcgacccct   20400
acttcggcgt gatcaccaac ccgcgcttcc gccgcgggca gatggacgag gaagcgaagc   20460
gcgaattcct cgcgtccggc cgcgtccatg ccgactacgt gatgcgcatg gtgcacgcgc   20520
acatcgcgcc cgacttccgt ccgcgcacca tcctcgattt cggctgcggc gtgggccgcc   20580
tcgtcatccc cttcgccgcc caggccgagc aggtgaccgg cgccgacgtc tccccctcga   20640
tgctggccga ggctgcgcgc aactgcgccg agcagggcgt ggccaatgcg cgcctcgtgg   20700
tctcggacga ctcgctcacc gggctgccgg gccccttcga cctcgtgcac tcgttcatcg   20760
tgttccagca catcgatccg gcgcgcgggc gcgacatctt ccgccgcctg ctgggcaccc   20820
tggcgagcgg cggcgtgacg gcgttgcact tcgtctatgc gaagcgtatc tacgcggcgc   20880
cgtacggcgt ggcgccaccg cccgagcgc cgccgcccc gccgcgcccg cagcctccac   20940
ccagccgggc ggagatcaag gccgcgagca aggcgcgggc cgcgctggcc acgcaggcac   21000
gggcgccgga accggcaccc aatcccgacc ccgacatgca gatgaattcc tacccggcgg   21060
cggagatgct gttcctcgtg caggaagccg gcgtcacgcg cttccacgtc gagttcaccg   21120
accacggcgg cgagctgggt ctcttcctct tcttccgcaa gccctaggcg tccttcgagg   21180
cgaaggagag gcccagcgcg gtgatcgtcg cggcgtgctg cggggcgacc gccgccgcgg   21240
ccgcgacttc ggcagccacg cgctcgaggt cgccgcgcga cccgtacagg cgcacgcgcc   21300
agtagtgcac ccagtaggag tccgacagg ccgacgcgat tccagcgcg cgctcgacct   21360
cgggcccgac gcgccctgg tgcacttggt gcgaggccca ttgcaggttc acggagggct   21420
cgaggggcag gacggcgagc agcgcctcgc tcaccgccag cgcgcgcgcc tcatcgccgt   21480
cgtcggcgag gcgcatcagc gtgggcaggg actccggcgc aagcgcgtgc gcggccgcga   21540
gcgaatcgcg cagcatcgca gcgtcgccgc gcgcgagcgc caggcgggcg cggtggtagt   21600
gggtccagaa cgaaggcgcc gcggcatcga gcgcgcgcgc gacctcggcg cttgcgtcgc   21660
ccccggacag ccgggacgcg gcccacaaca ggttcgcctc cgccgcgtgc gccggcacgc   21720
tccgcaacgg cgccagcacc gcggtggcga cctcgcgcga accctcgcgc gacgagaagg   21780
ccgagcgcgc gaactgcatc agcgattcgc cgagcgacga atgttcggcg cagagccgcg   21840
cgggatccac cgcgccgagc gcgcggcgg cacgcagcgc cgcctcctcg tcaccgcgaa   21900
gcgcgtgcag ccgcaggcga tggaacgcgg cccactccgc cgacatcgac gccgccccgc   21960
gctccagcgc acccctcgacg gcgccatcgg tgcgcccgcg cgcgatgtgc aggatggccc   22020
agtcgcctg cgcgccgggg aggttcgtgc cttgcgcgac gagcggtcg aggaccgcga   22080
gcgccgcggc ggcgtcgcgc gcgccactga agtcgttggc caggccgatc agcgccccgg   22140
```

```
cgagctcggg ggccgcgcgg aaggcggccg ccgcgtcgcg ccccaccgcc gcgaggtcgc    22200
cgcgcagccg gtagaggcgc gcgcggtggt aatagcccca gaacgggtcg gggctcgcga    22260
tcagtgaatc gaacgagcgc tcgacctcgc cgcccacgtt gccgaggtcc acctggtgcg    22320
cgccccacag ggtcgccgca tgcgcggcca tcggctcgtc ggggacgaag gcgcggtgga    22380
tcagctcgac gtcgccgaag cgccgctcgc gcaacgcgag ctccagggcg ttcagccgct    22440
cgaggctcgc cttgtactcg cgataggcat cgcgtcgcg gtcgtattcc tgcgcggcac    22500
tgtcgtccat cagcgaggcc gtgagccgct ggtcgccgtg aacggtgtcg tatgcgtagg    22560
ccgcgcgatc gtcgtagggc agcgtggccg cgcgcagcga cccgaacgcg gcgcggtcgc    22620
ggaatcccag cttgcccagc acgcggggtg ccaggtcgtg ggcttccatc tcgcccgggc    22680
tccagtcgcg cagcgccggc acgttcgagt agagccacaa cggcgtgcgg tagagctgcg    22740
ccaggtccgc gtccgggcgg ccctgcgcct cgaagatcgc cttccactcc gcgagtccct    22800
cgtggccggc ggagaagatc ggcacgtggt cgccgaagat cgccagcacg tagggccgcc    22860
cgcgcacctc gagacgctcg accagcgagc gcagcgcgcg gtcggtttcg gccatcaccg    22920
cgcagtactt gtcgatcgcg ccctggcgca cgtgcccggc gaggaagtgc ggcgccacgt    22980
agcgcggctc ggcgcccttg tcgcccaggt agggtccgtg gcccatcagc gagacgccga    23040
acacgaagcg cggccgcgaa tcgtcgagct ccgcgatcac ggcatccatc acggccgaat    23100
ccgcgatgcg ctcgaactgg cgcggcgcat gggcgaacgc ctcgtcggcg tggaacgccg    23160
cgaagccgag cttcttgtac acgcggtcgc ggttgtagaa ccagcggtgg aacggatgga    23220
tggccacgct gcggtagccg tggcggcgca gcatgccggg caccgcatac acccggtcgc    23280
gcaactgccc catgtagggc agcgcgcggc agccttcctc gcccaatgcg atgccggtga    23340
ggatctcgaa ctcggtgtcg caggtgttgc cgccgatcac cttcacgcgc aggcggccgg    23400
tggtgcggcc cggccgcgtc gcgccgagga acggcagcgg gtcgcgcgcg agcggcacgc    23460
cgaacgggcg caggtcggcg agcgactcca tcatcaggaa gatgatgtcg gggcgctccg    23520
cggtggccgg gccgctcgcc gagcgctcct ggcgctccag cagccgctcc acgccctcgc    23580
gcagccgccg cggcgcgacc gggcgggcga tgctcccggc gaccagcacg gcgtgcagga    23640
cgaagcccag gcgccgcacg ttcacgttgt ggtcggtggg cgaatacgtc atcgcgtcga    23700
actccaccgg ccagcggcgc aagcgtccgg cgacgaaggc gacgccgagc gcggcgagca    23760
cgagcagcca ctgccggcga tcgtcgtacg cgtcgtgcag gtcgatggcc tcggtgatct    23820
cgaacgcggc ccacgccagc gccgcgagca ctgcggcagc ggcaagcgcg cgcaccgcca    23880
cgcgcgggcc atgcaggtac gccggcgcga acaccagcgc ctcgcgagcg ctcaggaaat    23940
cgcccgggcc gaccacgcgg ccgaagaacg cgagcaggcc tgcgttcacc gcgtagagcg    24000
cgcccgccag cgccacggtc cccaccgtcg cgagcgcccc gaccgcgacg aagagcagcg    24060
catggaagat ggcgagcagg ccggcgttcc aggccaggcg ccgcgccgac gccacgggtt    24120
ccacgtcata gcccaccacg aggcgccgct cgagggcggc ctgcagcacc gcgtaggcga    24180
gcacgcagac cagcgtgccg cccagcagca cgccgaggaa acccgccaaa ccagccgtgt    24240
cgtcgaggat catgggacca ccagggacgt cttccggcca gtctataacg ccccgcccca    24300
tgaaaaacgg cgaggcccgg gggcctcgcc gtggagggaa ccgcgcaacc tcagggttgc    24360
ggggggtgcga ccgaccagcc gccggggccg atcaggtttg cgctgccgat ggtggcgagc    24420
ccgttcatca ggcgcaccgt gacgctgccg tcggtgtgcc ggaagaccag gtccttgcgg    24480
ctgtcgccgt tcaggtccag cagctgcgtc acgttccaac ccgtgcccgc cggcagcgca    24540
tcgcccgcgg tgatgatgga aacaccatcc atcaggcgga tgtgcacgcg gccgtcggtg    24600
tggcggaaga ccaggtcgtc cttcgcgtcg ccgttcaggt cgcccaccag gttcaccgac    24660
cacccgctgc cggccggcag gagctgggcg ctggcgccga aggtggtgcc gttcatcagg    24720
aacaggtgcg cgcggccgtc ggtgtggcgg aagatcatgt ccgcgcgtcc gtcgccgttg    24780
aggtcgccca ggtggcttac cgtccagccg ctggcggccg agaggaagcc cgctccggcc    24840
gtcaccgtgg tgccgtccat gatgtagatg tagccgcggc cgtcgctgtg catgaagacc    24900
aggtccgcct tgccgtccac gttcatgtcc cccgcctggg tgagcgtcca gccggtcccc    24960
gccccgaaga gctgggcgct gccgatgatc gtcgtgccgt ccatcaacca gatgtgcgcg    25020
cgtccgtcgg tgtggcgcag gatgatgtcc gccttgccgt cgccgttcag gtccgccgtg    25080
tggctcacga tccagccgag cccggcgggc agcagctcct tgccggccgt caccgtagtc    25140
ccgttcatga tgtacgcgta ggcgcggcca tcggtgtggc ggaacaggat gtcggcgcgc    25200
aggtcgccat tgaggtcggc cgtctggctc accgtccagc ccgcgccggc gccgatcagg    25260
ttggccgagc cggtgatggc ggtgccgttc atcgtccaca ccgcgatgcg cccgtcgctg    25320
ttctgcagca cgaggtcgct cctgccgtcg ccgctcaggt ccgagacgac cggcttggtg    25380
ggcggcgtga cgccaccgga aaccggcagt gccagcccgt cgatccaggc gcggtcggcg    25440
cagttggcgc catccggggc cggcgggttc caggtcggcg agttgcaagg cgtggagagg    25500
aagttcgaga aacgccacac cagcgtgtgc gcgcccgcgg tcaccgggaa cgacaccagc    25560
```

WO 2004/013327   PCT/EP2003/007765

31/39

```
ttccagcccg aggccgtggt gcccgcgtcg ctgaacacca ccgtgccgtc gatcaggaac    25620
tcgaacttgc cggcgttggg gaagctcgac acgcgatagg cgaacgcgac gttgcccgcg    25680
agcagcgtgc ccgcgaacga gaggtcggag ttcaccgtcg tcgagttggt cgggtcgctc    25740
gtcaccacct gcgccgaacg aaggctggtc gcgccctcga aggcctggtc ggagccgacc    25800
gtccatgcgt tggcgccgcc cgaggtcgtg aagccggcag gcagcgtgcc acccgtgggc    25860
cacggcgaga ggatgatgcc ggtcgcggtc gccgggccgc cgatgaagac acccgtggaa    25920
cccgtcgcat tggagagcgc gacgttgaac gtctcgttgc cttcaccgat accgtcgctc    25980
gccaccggga ccgtgatggt cttgggcgtg gtctcgccgt tggcccagct gagcgagccc    26040
gaggtcgcgg tgtagtcggc gcccgaagtg gcggtgccgt tggaggtggc gtagctgacc    26100
gagatggcac ccgccgagcc gccgatgcgg ctcacggtga gcgtcacgtt gccggcggtt    26160
tccgcggccg cgaaggtggt gccggtgaac tgcaccgaac cgggaacggc gggcgggggtg    26220
accgtggact tgagcgccga gagtgccgcg cggttgttgt tcagcgagag cgcgtcgttg    26280
gcctgcgcgg tgccgcacgg ttgcggcgca ccgacacccg tcgtgcacga gaggcccggg    26340
ttggagaagc ggtagacctt ggtcgtcgag tcatgcagca gcgacatgat ggtggtgaag    26400
ttcgccgcgt cgttcgtcgt gcactcgggg cgctggctca cgccgcaggc gccgccgccg    26460
gagccgttgg gcagggccga gttgcaggtg aggttgccgc tcgcgcaata gaagtggccg    26520
tacgagtacg cgaacgcacc ttgcggcggg ctcgcggtgc cgcccgcttc ccaggcctgc    26580
gaggggcggt cgtggcggtt gcccatggca tggccgagct cgtggatgaa gacccactcg    26640
caaccgaaca ggcagccgga agcaacggcg taggcaaagt tggggtccgg cgtgctggtg    26700
ccgatccagg ctgccgccgaag tcggcgccgt tgcgcatgaa ggccaccatg    26760
tccgcgccgt attgcgtgcg gatggcctcg atgttgccga acgtcgcggc gtcgaagctc    26820
gcgtcgcccg gggtgatggc atgcatcgcc gtggtgtcgt cgatggcgtc gtcgtagttc    26880
acctgcgtcg cattcaccag gcgcagcgtg atggcgacct cgctgtcggc atacgcagtg    26940
ttggcgcgcg tgacgaggaa gttcaggcgc gtcatcaggc tgccgccgag gcgctggccg    27000
aagccgcgcg agtacacgat catcaggtcg atgacgttct gcggcgtggg cgtcacctgg    27060
gcgagccgac ggtattcacg ccggggatgg cgatcagggt ctcggggctg ccggtcgcga    27120
cgacgcgcga cttcgagccg gtcgtggcca cggcgtggtc ggcgccgagg ctgatgatcg    27180
gcatgctggc ctgctcggcc gtcatgtcga ccaggtagtc gacggtgccg cccggaatca    27240
ggcggtactc accctgcggg gtcgagaaca cgccgaagga gccttcaggc ccggtcgtga    27300
cgatcgcgcg gtggttgatg ccgtcggtct tgctcttcgc gatccacgac gtgatgccgt    27360
cgccgtgggc ctggagcaac tcgaagacat acgagtggcg cacccgttg gggagcgaca    27420
gctcgacctc ggagcgcgga ctcagggcat ggagggccgc ggcgttgaag cgcaccgctt    27480
gccgggcaac ggcgctgccc ggcgccgaca tcgcgggcgc gggtgcggag accaggatct    27540
ggggaaccgc ggcgaaggcc acggaagaga agacgatcgc cagcgcgccg aggaaggcgg    27600
ctgtgatccg ttgcgtgaat gcgttcatgt ggagctccgg aagttgaccc atgcccaatc    27660
cgctatgtcg cggagatgtg gacaaaaggt atcgagcggg cgtgacgacc cgccccgga    27720
gggatgctcc aaaaggacta cgagggtgct acggctgggt tggtggcgaa ggccgaagga    27780
gaactccttg tggtctccgc gcgacttaag gttgcgaggg aaccaccgac cacccaccgg    27840
ggccgatcaa gtccgcgctg ccgatcgtgg tgacgccgtt catcaggcga atcgtcacgc    27900
gcccatcgtt gtggtggaag accacgtcca tgcgcccgtc gccgttgaag tcgaggagct    27960
gcgtcaccgt ccagcccgcg aacggcggca ggatgtcggc ggcgccgagg atcgcggtgc    28020
cgttcatcag gcgcacgtgc gcgcggccgt cggtgtgcgc gaacacgagg tccgcccggc    28080
cgtcgccgtt gagatctccc accaggttca ccgaccatcc ggtgcccgcg ggcaacagct    28140
cggcgctggc gccgaacgtc gtgccatcca tcaggaacag gtgtgcgcgg ccgtcgttgt    28200
ggcggaagac gaggtcggcc ttctggtcgc cgctgaagtc cgcgacgtgg ctcaccgcca    28260
agccgctgcc gggcgcgacg aagcccaacc cggcggtgat cgccgtgccg ttcatgatgt    28320
atgcgtagcc gcgcccatcg acgttggaga agacgatgtc ggccttgccg tcgccgttca    28380
tgtcgccggt gccgaccacg ttccatcccg agccgcgggg cagcaggctc gcactgccgc    28440
tgatcgcggt gccattcatc agccagatgt gggcgcgtcc gtccgtgtgc tgagcagcag    28500
gtcggccttg ccgtcgccgt tcaggtccgc cgtgtggctg atggtccagc ccgtgccggc    28560
ggggaacagc tccttgccgc ccacgaccgt gagcccgttc atctggtacg tgtagatgcg    28620
cccgtcggtg tgctggaagt agatgtcggc ctggctgtcg ccgttggaat ccgcggccat    28680
tgatcacgct ccagccggcg ccggcgggaa tcaggttcgc cgaaccagtg atcgtcgttc    28740
cgttcatggt ccacgcggcg atgcgcccat cggtgtgctg gaacaggaag tcgctctgcc    28800
cgtcgccgtt caggtcggtc gccgcatggg tcgggaacgc caccttcgcg aggaaggcat    28860
cggagaaacc cgtgagcaac gtcttgtacg cgccgggcgt ggtcggatag ttcgccgacg    28920
acgtgcggcc ggccacatag gccgcgcccc gcgcatcgac gccgacgccg tacccacct    28980
```

```
cgacgttgtt gccgccgatc agcgtcgaga acacgatgtc gccgttgcgc acgcgagtga   29040
ggaacgcgtc ggcatcgccg ccggggccct gcaccgggtt caggtcgggg aagccggcga   29100
cctgcgtgta gcccacgacg agcgcatcgc cattcggcgc cagggcgacg tccaggggat   29160
tgtcgtaggc cgagccgccg aagatgcccg acgattcgag cgcgccggtg gccgagaagc   29220
gcgtcacgaa cgcatcgacg gtgccggcga aattgcgcac cgttccggtc tgcgggaagc   29280
tgggcaggcc ggtatcgccg acgacggtcg cctggcccgt gaccgggttc accgcgatcg   29340
cggtcacgtt gtcgccgctg ccgccgagga acgtggaata cgcgaccggg ccgccgctcg   29400
caccgatctt ggtgacgaag ccgccaccgg gttgggtctg gtacgcattc acgttgcgga   29460
acgtcgagga gcgcgtggag cccgcgacgt acgtgttgcc cgcattgtcc gttgcgatac   29520
cgcggccgat gtcatcaccg ctgccgccga gatacgtgga atagacgaag ccgccgttga   29580
acgagagctg gtgacgaacg cgtcgcgatt gaggccggtg ccgacgaaca cggctgcagc   29640
gcgctgccgc cggtcgggaa gccgcgcgcg gccgaacccg tgacatgcat gaagccgttg   29700
ttgtccacgg cgatcgccat gccctcgtcg tcgccggagc cgcccagcag ccgcgagaag   29760
atgacgttgc cagtggtatt cagcgccgtg acgaatgcat cgcgcccgcc ggcatggttg   29820
gtgcaatcgg tgacggcgaa gcaatgcggg aacgcgttgg tgccggtgga accggtgatc   29880
gcgacgaact cgttgccgcc ggaggcgccg agcgccagcc cgcggagata cgcttcggcc   29940
ccgagcgtgg tcgagtacac cacctgctgc gtggtggggt tgaacttcac caccgtggct   30000
tcgtacgtgc ctgcggtata ggtcgcgagc ccgacgtaca cgttgccctg cgcgtccacc   30060
ttgaccggcg tggtccactg ctcgtcgccg gcgccgccca ggtacgtgga gaacgccatc   30120
accgggtcga tcaccagcgt gcgggtgctg tcgtagtcgg cgatcacgaa gccggcctcg   30180
gcgccctgct cgcccacgaa gagctcgaag cgcgcggcca ccggcacgcg cgaatcgccg   30240
acctgctgga aggctaccgg cgcgtgctgc gtaaattcct cctggcccac gcgcatgcgc   30300
acgttgccct cgccatcgat ccacgcgtgg tcggcggacg agaggtcgag gcggatctgg   30360
cgcggatcgg cgcgcggggc gaccacgaag tcgtattcga gcgtgccctc cttgccgtac   30420
acggtgaggt ccacgccgcg atacaggtcc ttcagcgaca cgcggccgaa gtgcggcacg   30480
ttttcgcgcc ggccggccggc cgtggtgccg ctgtaatagt ggctgagggt ggcccggggc   30540
tcctcggcct cgatgaccgg atcggaagcc cccgcgaagc gcacgcgcaa cagcgaggcg   30600
ctggcagctt tcctggagcg tggctcgaag gcaatgccat cgcggctcac ggagacgcgt   30660
ccggcctgcc cgcgcgagac gtagagcgca tccgggccga actggccctc gttgcgctcg   30720
aacgtgatgg gcgcgttcct gagggcatca gggacgacgg catggacggc ggcgggagtg   30780
gcgagggcgg ccgcgacgag cacggccagg gggcggagcg ccgggcgggc gccgctggag   30840
ctgcgggtca tgggcagtg accttctgt tgttatggct gcttgttgtc tttgcgggat   30900
ccacagctcg gatccagctg gcgatactac cagcgagcga gtggtttttc gtgacccagt   30960
cccgcagtcg ggccccggcg acaacggcgg catcgcggtc ctttgccagg gcacgcacgg   31020
cctcgatcca gtcgccgtcg cgggccagca cggcgggggc atcgcggtag ggctcgagat   31080
ccgaggccac caccgggatg ccgaggcagc cgtactcgag cagcttcagg ttgctcttcg   31140
cgcgattgaa cgggttgtcg cgcagcggtg ccaccgcgac gtcgagtgcc agcgacgcga   31200
gcttctccgg gtactgcgcg atgggcacca tgtcgtgcac ttccgcggcg aacggcgcca   31260
gctcgggcgt gcacaggccg aggaacaccc agtcgatctc gcgatgcgtg gcgcgcacca   31320
cgggctcgag caacttcagg tcttccccat gctgtttcgc ccccgcccag cccacgcgcg   31380
gacgcgcgcc acccgcggga cggttcgcga ggccggccca acgctccgca tcgatcgcgt   31440
tcgggatcac gcgcacgtcc ttcgcgccgc ggccgaaggc ctccgcgagc ggtgcggtgg   31500
acaccacgag gcggtcgcac aacgccaccg cgcgcgcgat gcgctgggcg atgtccgggt   31560
agatcgtcgc cgcatacgga ttgccgggcg gcagttgcgt gagcaggtca tccaggccca   31620
ggaccttcag cgcgcgtccg tggcgcgcga gcacctcgag cgaggtgagc tggtagttgt   31680
ggaagaagtt gtgcgcgagc acggcgtcgg cgtcgaggcg ctgccactcg acgcggttgg   31740
gcgcgcagcc cgtggcgtgc tcgcccatca tcacgacctc gacgcggccc gcgcgctcca   31800
gcgccgcgca cggctggcgc acgcgcacct cgccggagcc ccagcggtcg aacggaaacg   31860
cgcagagccg aaccgcgtgc gtgtcgttgc cgcgcggcgc gaagcgttcc acgggcccgt   31920
agccctcgcc gcccatgcgc aacgccgggt ggtagtgcgg atcgtcgtcg agcaccggct   31980
gccagcgctc gcgcatccac gcctcttcgg aaggaatgac gagcttcgcc accggaacga   32040
cggcatcggc gcgcgcatcg gcgagcaacg ggaagtccgc cgcgatgccg gggcgcgcaa   32100
ggatgtcgaa gcccgctgcg cgcagcccga ggcacaggtg cgccatcgcg aacggccccg   32160
cgcgctcgat ctcgtgcgac gcctgacggc agagcgccgc gtcgcggttg accagcgcga   32220
gacggcccgc cgcgacgctc acctcgcgcg cgatgcgta gagtgccgcg aggcgctcct   32280
cgccctgcgc gcgcggcgga gcgcccgaga cggcccacgg cccgccgccc gcgatttccc   32340
agccgggaac gcgcacaccg ccggggccca gcaggtcggg cgcgagggca cccgtggagg   32400
```

```
ggcctgcgat gccctgctcg aggcgctcga gccagccggc gcggaattcg ctgcaacgcg     32460
catcgacgat cgcgatccac agcgtcgccg cgttgcggat ggcgcgcgcg aacgcggagc     32520
cggtttcgcc gcccaccgga accaggtgcg tgcgcacgcc gcgcccgccc agctcaccgg     32580
gccagcgttt cggcgccctt cgcgtcgaga tcgacgtaca cgtggccgat acgttcgcgg     32640
cagcccttca cgagcgcctc gatgcacgcc accagccgct ccggcgcgtg cgtcgcgtca     32700
atgaaggcgc aaatcgtgcc ggccccgatc gcgcgcgccc acgatgcggg cgcatcggac     32760
gcagctggca ggaacgcggc ctgcgtgcgc gtgcggcgga agtgctccgc gaccacgcgc     32820
tggtggcgct cgcgcgcggc gtcgctttcc tgcgcggcga tgttcaggtc gaggcgatgc     32880
acgagcagct cgcgcagtgc caccacggcg ctcgcacggc cctcctccgc gaggcgcagc     32940
gcgagatcgg tgacgcccgc ccagccgccg gcgcgcagcc cgccacgctc cagcaacgca     33000
gcgcggcgca gcaccgacga cccggcgagc ggctggtgtc gcgcgaactc cgcgccgaac     33060
gtggggacga acgccggatc gtagcgttcg ccgtgcgcgc tcagggcatc actgcagccg     33120
tagagaagct cggcgttcgg gttgcgtgcg accgcacccg ccaccgcgct cggcgcgcca     33180
tcggccagcc ggtcgccctc ctccaccagc acgacccatt cgcccttcgc gtccgcgatc     33240
gcgcggttga cgcgcccgag cgccgcgtcg ctcgcgtcgc gctcgaggac cacggtctcc     33300
agggcgatcg cggcagtgct cctggccggc aaggcctcgc gcgccatcca gtggtcgtag     33360
gcatcgccgc ccagcggcgt gcgctcgcgg cggcgtggaa agccgaggca cgccgcggcg     33420
aacaacggat cgatcgttgc gggcacagca tccaccgcgc catcgcgctc cgccgccagc     33480
tcgcgataga gcgccaggta ctcggccgca ttcgcctcga tcgtcttcgg gcggtggtac     33540
ttcaggttct ccgcgacgcg atcgagcgcc gtgcggtcgt cgcgcaggcg cagcaggagg     33600
gccacgatcg cgccgggatc gttgtgcggc agcaggaagc cggtctcgcc atcgcgcacg     33660
cgctcgggaa tggcgccatc gcgcgtgccc gccaccggga tgccgaggcg ctgtgcctcg     33720
ctcaacacca ggcagaaggt ctcctccatc gtcgagggca gcaggaccag gtcgatgccg     33780
cgaaggatgc gcggcaggtc cgtcgcggca taggcgccgc gctgcatgat gcccgccgca     33840
ttcgccgccg cttcgagcgc gagatcgcgc ggaccccagg cctcgatcgc gatgcgctgc     33900
cccgccagcc ggcgcgcggc ctcgatgatc aggtgcgcgc ccttgcgccc gccgaagcgg     33960
ccgaggaagg ccacgcgcag caccggccgc gcttcggggc gcggaagcag cggcaggtcg     34020
gcaatgccgt gcccgatgac gcgcagcttc gcttcgcagc tcgcgccgaa gcccttgcgc     34080
atgcgcatcg ccgcgtactc ggacgggcag aggatcacgt cggccgcgtc cacggcatcg     34140
gacgccgcgg cgaagcgcga ttcgaggtag ccggtcacgg acgccggcgc cgtcggcccc     34200
acgatcgcct cgctgcggcc gcgcaggcaa tccacgcatc cgccatcggc gccgcgcgcg     34260
gcggcgcggc cgcacggcag gtcctgcggc ccggtcatca tgttgtagtc ggcgcacagg     34320
tacgagaagt cgtgcgccgt gatgaccacg cgtgcccccg actcacgcgc gatgcgcggc     34380
aggcgcaggc tgttccagcc caccaacgac tggaagtgca cgacgtcgta gccccccggac     34440
agcagttcgc ggaaggccgc ctccacgccc gcgtcggcaa ccgccgcctc gtgcccgatc     34500
acgcggatgc cgggcgtcat gcgctcgggg ttcaggcgca cgagtcgcac gtgcgtcgcc     34560
aggcgctcgg tggtgaagtc ggtccacgcg cgcgggccct cgggcagcac cagcgtcgag     34620
gcgacgtcct tgcgcagcag gttcaccagc gcgcgcgtgt gctcctcaat gccgccgcgc     34680
gagtcgaggc gatgcagcac ctgcagcacg cgcgggcggc ccggcatccg ctcgcgctcg     34740
gccgcggcgt tgacgcgctc ggcggcgcgc cgcagcgggt tcgcgatcgc ccacgcgcgc     34800
acgcccgcga cgtaggtggg ccagcggcgc tcgaggcgca cgcggttggc ctcgcgcgcg     34860
gcgtcgatgc cggccacccc gccgaacgag gcctcgcccg cgtggtgcac gtagacgtcg     34920
tcgcagcacg cgatctgcca gtcccggtcc agcgcgcgca tcgagaggtc gttttcctcg     34980
ccgtagccgc ggccatagag cgagtcgaac gggccgcatt gcacccacac ctcgcggcgc     35040
accagcatgc agaagccgac cgccgtcggc aggcgcggat actcgcggcg gctggcgcgc     35100
gcgacgcagg cggagatggc atcgacgtcc ggctcgccgt cgggccgctc gaagatatcg     35160
gcgacgccgg gcaggctgag gatcgtcgcg cggttggaga gcgggcacac gatgccgacc     35220
cgcgcatcgc tgtcgcggca acgcagcagg ccctcgagcc aaccggggcc gacctgcgta     35280
tccgaattga gcaccacgaa gtcgccgccg ctgccggcca ggccgcagc gaccgcgcct     35340
gcgaagccga ggttcttcgc attggtgacg acacgagtcgt cggcacgacg cgccgcgaag     35400
ctgcccagca acccggcgag gcgcggatcg gtgctcgcgt catcgacgag gacgatgccg     35460
tgcgcggcga ccgtgttggcg ggccagcgac gcgaggcagc gttcggtgtc gtcgtggccg     35520
ttgaagaccg ggacgacgac tagcgccggc gccatccgac cggcgcttcg agctgggcca     35580
atcgcgccag caggcgctgg tgctcctgcg agagtgcctc gagctgcttc tgcgaggcga     35640
cgtaggcgcc tccatcgcgc cccaggtgcg tgcgcgcctc ctcgagctgc cgcagcgact     35700
gctcgtggcg accggagacg tcgtggctcgc ggtcggagag cagcgagaag cgcccggccc     35760
ccgccgtcac cgcctcgcgc tcgccgcaca gggcgaggaa gtacatcgca tcggcgacgc     35820
```

```
ccgccgcggg tttcgcagcc tccggcgccg aagcttgcag caactgcacg gcgtcgggca   35880
ccgcgtgcag cggccagatc gccgaatacg catcgacgcg ttgcccgaac agggccacct   35940
ccgggaaccg cgccttcagc acggcgagga attcgtgctc gtagagctcg cgcacgtgat   36000
gcgggttgcg gtagtcgcgc tggtccgagt acacctcgcg gttcggcgtc gagaccagca   36060
gcaggccgcc gggcgccagc acgcgcttcg cttcgtcgag caggcgctcg ggatcgggaa   36120
tgtgctcgag cgtctcgaag gagacgagga gatcgatgct tgcgtcatcg cacggcagcg   36180
actcgcagcg gccctcgacg tactgcaggt tctgcgcggc accatagcgg cggcgcgcct   36240
gcgcgatggt ctcggcggcg acatccgcgc ccaccacgga cttcgcgcgc gtcgccagca   36300
gcgccgagcc gtagccctcc ccgcaggcga tgtcgaggac gcggcagcct ccggcgagcg   36360
gcaacgcaaa gtggtagcgg tgccagtgct cgtaccagat ctcgccctcg aagccgggct   36420
ggaaacgttc gtgttccatc gggagagtat aggtggatgt gcaaccgccc tccggagagg   36480
gcggttgcga tgatcttatt cgagatcggc tgctacggaa gcaccggcaa cggctgcggc   36540
ggcgccaccg accatccggt gcccgcgccc agcaggttcg cgctgccgag cgtggtgagg   36600
ccgtccatca ggcgcaccgt gatgcggcca tcgacattct tgaacaccat gtcgagcttg   36660
ccgtcgccat cgaagtcgag cagttgcgtg acggtccagc ccgcccccgc cggcagcacg   36720
tccgccgcgc tgaggatggc ggtgccgttc atcaggcgca cgtgcgcacg gccgtcggtg   36780
tggcggaaca cgatgtcacc gcgcccgtcg cggttcatgt cgcccacgtg gctcacgatc   36840
cagccggtgc ccgcgccgag gagctcggaa ccggcgccga acgcggtgcc gttcatgatg   36900
aagagatgcg cgcggccatc ggtgtggcgg aagaacatgt ccgccttgcc gtcgccgctc   36960
acgtcgccca ggtgcgtcac cgtccagccg ctggcgggcg agaggaagcc cgcgccggcg   37020
gtgatggtgg tgccgttcat caggtagatg tagccgcggc cgtcagcgtg gatgaagacg   37080
atgtcgtccc tgccgtcgcc gttgaggtcc ccggtggcca cgactctcca gcccgtcgcc   37140
ggccccagca gctgctggct gccgatgatg gccgtgccat ccatcagcca caggtgcgag   37200
cggccgtcgg cgttgcgcag caggagccgc gccttgccgt cgccgttcat gtcggcggtg   37260
cggtcgatgc tccagccgag gccggcgccc agcagctcct tgccgccggt caccgtgagg   37320
ccgttcatcg tgtacacgta cacgcgcccg tcggtgtggt gggatcctct agagtcgacc   37380
tgcaggcatg caagcttgag tattctatag tctcacctaa atagcttggc gtaatcatgg   37440
tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttcacacaac atacgagccg   37500
gaagcat                                                              37507
```
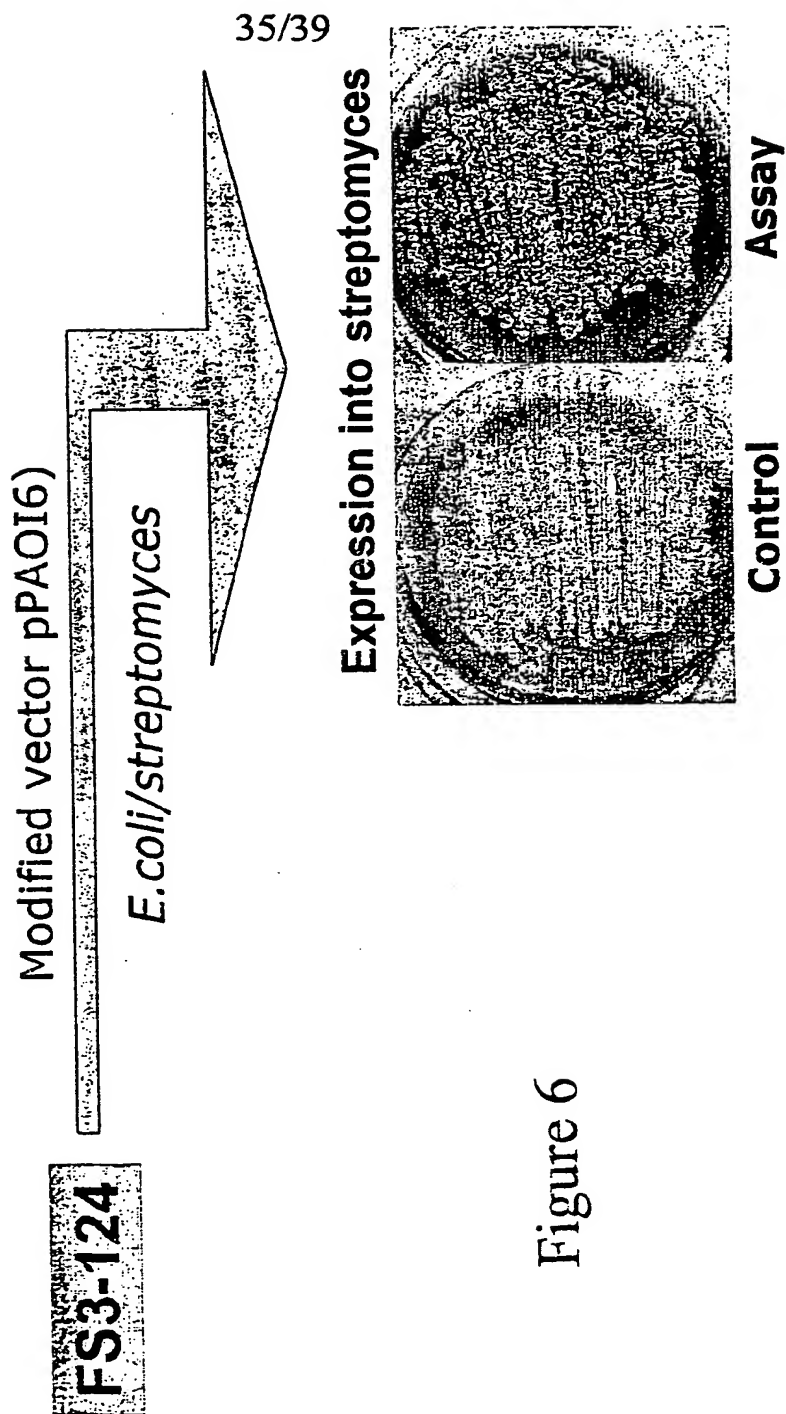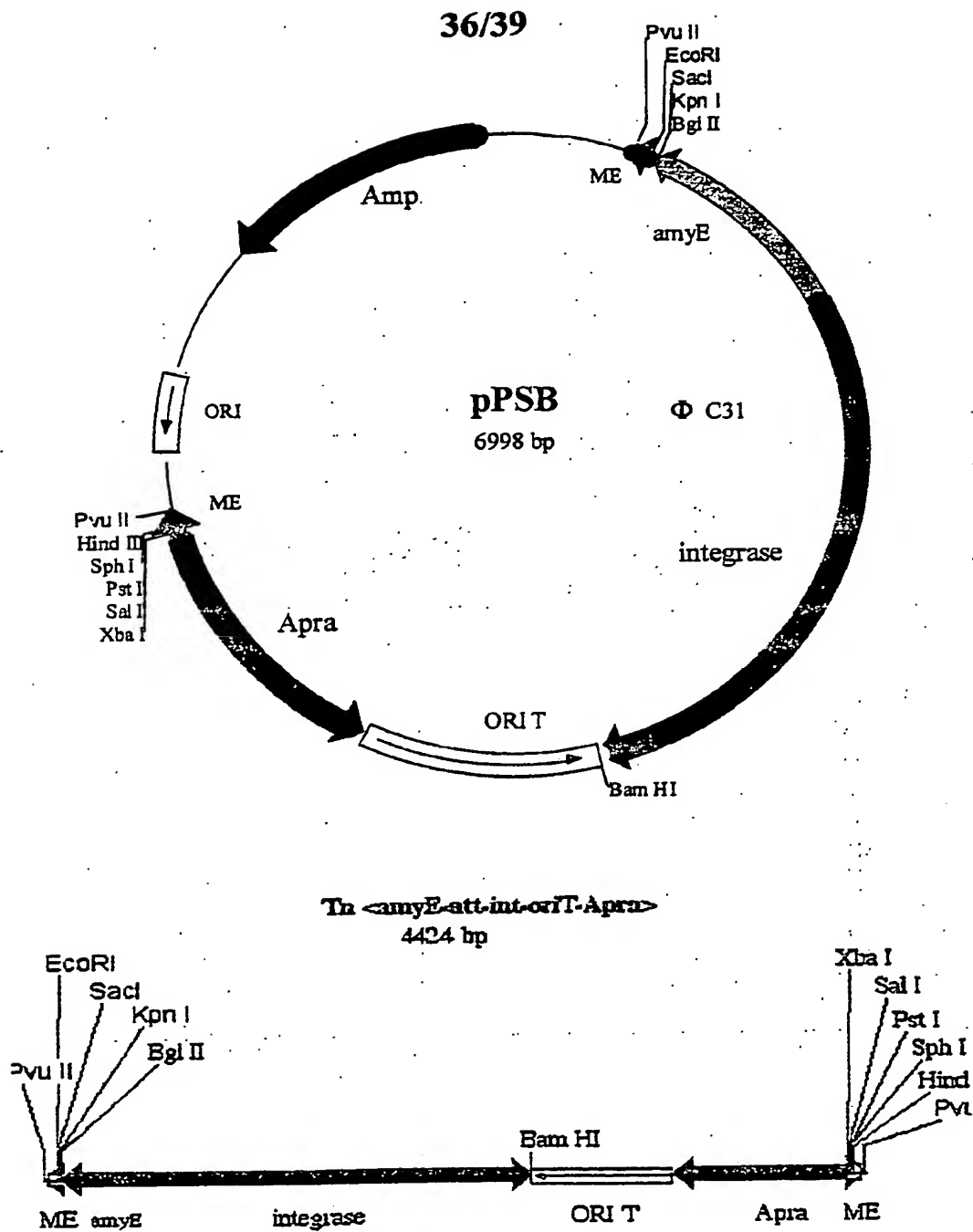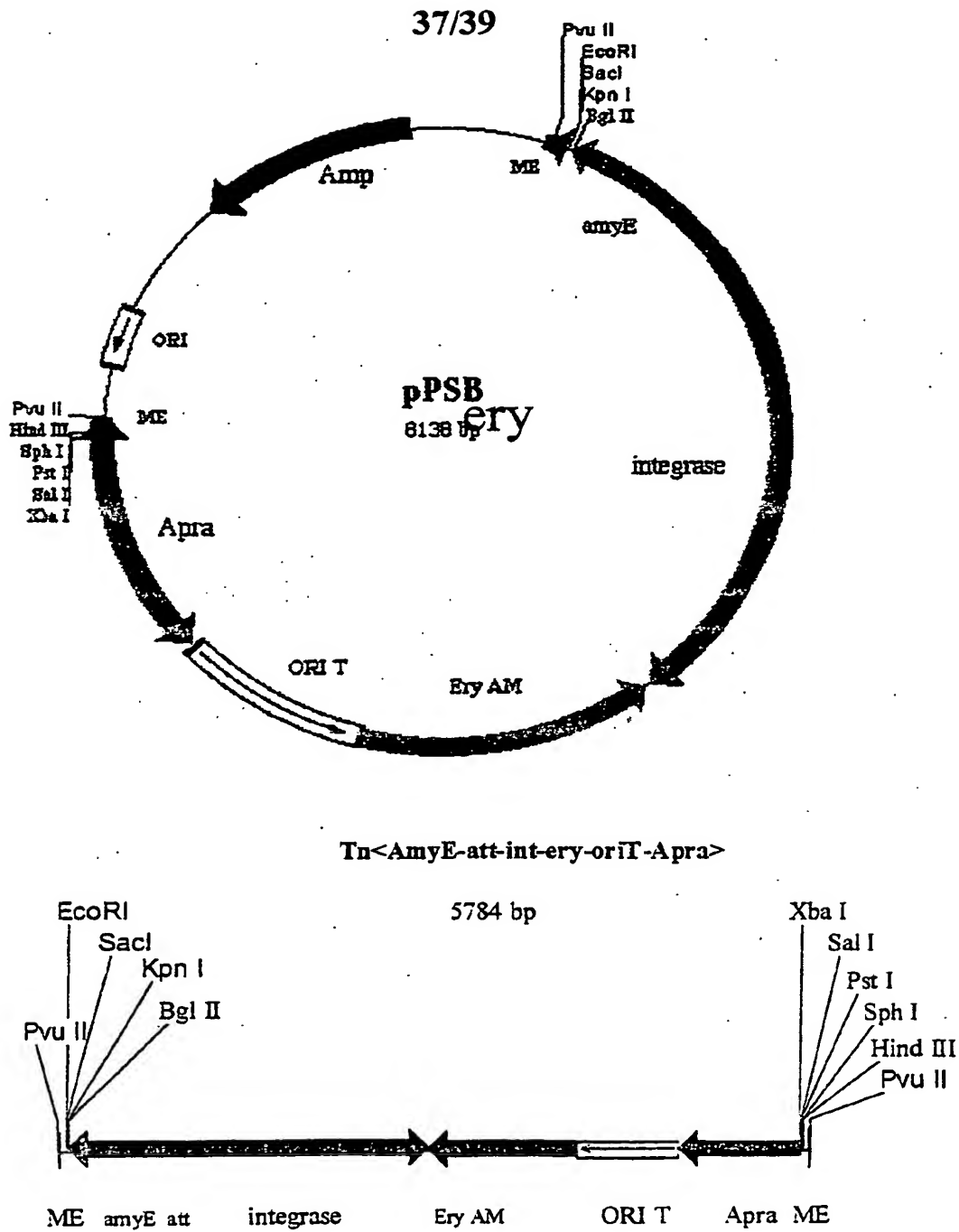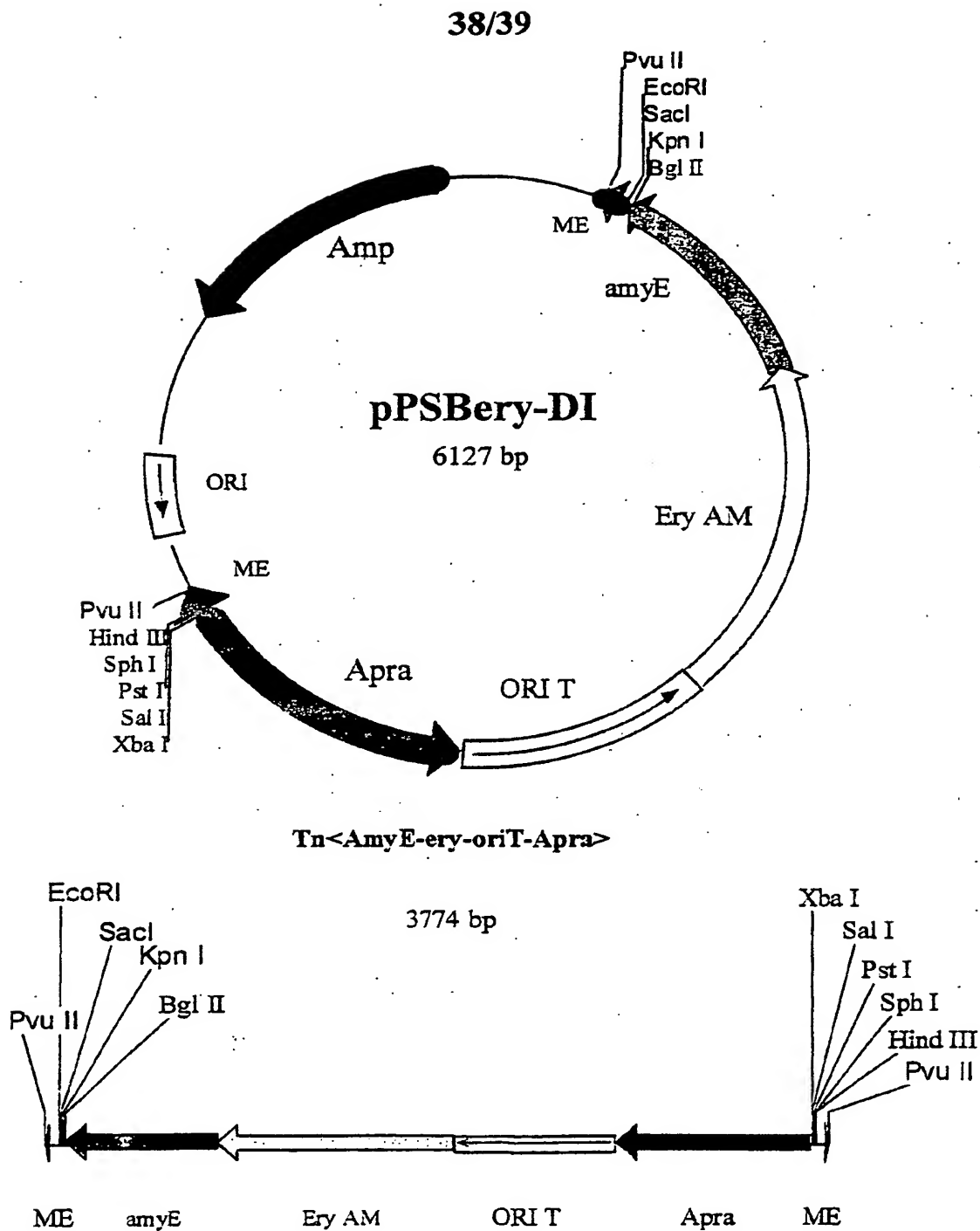
Figure 6

36/39

Pvu II
EcoRI
SacI
Kpn I
Bgl II

ME

amyE

Amp.

pPSB
6998 bp

Φ C31

ORI

ME

integrase

Pvu II
Hind III
Sph I
Pst I
Sal I
Xba I

Apra

ORI T

Bam HI

Tn <amyE-att-int-oriT-Apra>
4424 bp

EcoRI
SacI
Kpn I
Bgl II
Pvu II

Xba I
Sal I
Pst I
Sph I
Hind
Pvu

Bam HI

ME amyE   integrase   ORI T   Apra   ME

Figure 7

Figure 8

Figure 9

Figure 10